

Journal of Institutional Research South East Asia

JIRSEA-UPM Special Issue:
Vol. 24 No. 1

April 2026

ISSN 1675-6061

**JIRSEA is indexed with SCOPUS,
CABELLS Scholarly Analytics and
EBSCO Host (Education Research Index/
Education Research Complete), DOAJ.**

JIRSEA © SUIE 2026



JIRSEA EDITORIAL BOARD

CO-EDITORS

Professor Dr. Jang Wan Ko

Professor, Department of Education

President, Institute of Educational Research

Sungkyunkwan University, Seoul, Korea

President, Korean Association for Adult & Continuing Education

Email: jakosu@skku.edu

Associate Professor Dr. Teay Shawyun

Asst. President, Siam University, Thailand

Consultant, King Saud University, KSA

Southeast Asia Association for Institutional Research

e-mail: jerry182122@yahoo.com & teay@siam.edu

MEMBERS

Dr. John A. Muffo,

Director of Academic Assessment Emeritus,

Virginia Polytechnic Institute and State University, Blacksburg, Virginia,

USA. Contact: johnmuffo70@gmail.com

Dr. Sutee Sujitparapitaya

Associate Provost, Office of Institutional

Analytics, San Francisco State University,

California, USA. Contact:

sutee.sujitparapitaya@gmail.com

Professor Dr. Narend Baijnath

Chairman: Board of Governors, Commonwealth of

Learning, Research Associate: University of Johannesburg,

Member: Governing Council, Magna Charta Observatory,

Member: Academy of Science of South Africa.

Contact: narendbaijnath@msn.com

Professor Emeritus Edilberto C. de Jesús

Professor Emeritus, Asian Institute of Management, the Philippines,

Senior Research Fellow, Ateneo de Manila University School of Government,

Board of Trustees, Far Eastern University,

Board of Trustees, Phinma Education.

Contact: edcdejesus@gmail.com

Professor Dr. Sumalee Sungsi,
Professor, School of Educational Studies,
Sukhothai Thammathirat Open University,
Thailand. Contact: sumalee.sungsri@gmail.com

Assoc. Prof. Dr. Vimolwan Yukongdi
Faculty Director of MBA Program,
Asian Institute of Technology,
Thailand. e-mail: vyukongdi@ait.ac.th

Ir Ts Dr. Poh Kiat Ng
Deputy Dean (Academic & International
Relations), Faculty of Engineering & Technology,
Multimedia University Malaysia.
Contact: pkng@mmu.edu.my

Dr. Ng Gan Chee
Owner, Australasian Consultants, Melbourne, Australia.
Contact: gcng2001@yahoo.com.au

Dr. Raj Sharma
Senior Consultant in Higher Education, Melbourne,
Australia. Contact: dr.rajendra.sharma@gmail.com

Dr. Antonius Tanan
Educational Advisor, Jakarta, Indonesia,
Independent Commissioner at PT Ciputra Development & PT Metrodata Electronics
Tbk Contact: antonius.tanan@ciputra.ac.id

Professor Dr. Maria Pressentin
Professor, Strategy & Leadership, International School of Management, Paris,
France Chief Knowledge Officer, Leadership Scientific Pte. Ltd.
Contact: maria.pressentin@leadershipscientific.com

Professor Somwung Pitiyanuwat
Professor Emeritus, Distinguished Scholar in Education,
Chulalongkorn University Bangkok, Thailand.
Contact: somwung.P@chula.ac.th

Emeritus Professor George Gordon
Education Enhancement, University of Strathclyde,
Scotland. Contact: g.gordon@strath.ac.uk

JIRSEA Key Contacts (Effective 1 June 2025)

JIRSEA (Journal of Institutional Research South East Asia) © **JIRSEA 2002**
ISSN: 1675-6061

JIRSEA Publisher Address:

Sungkyunkwan University Institute of Education
25-2 Sungkyunkwan-ro, Jongno-gu, Seoul,
Sungkyunkwan University, Republic of Korea

Contact:

Paper submission and all queries:

Editorial Manager: jirsea@skku.edu

JIRSEA EDITORIAL and REVIEWERS TEAM

CO-EDITORS

Prof. Dr. Jang Wan Ko

Professor, Department of Education
President, Institute of Educational Research
Sungkyunkwan University, Seoul, Korea
President, Korean Association for Adult & Continuing Education
Email: jakosu@skku.edu

Assoc. Prof. Dr. Teay Shawyun

Asst. President, Siam University
Consultant, King Saud University
Interim President, Southeast Asian Association for Institutional Research
Email: jerry182122@yahoo.com; teay@siam.edu

REVIEW TEAM MEMBERS

Prof. Dr. Ma. Florecilla Cortes Cinches

Vice President for Research, Extension, Planning, and Innovation
Liceo de Cagayan University, Philippines
Telephone: +63 088 8584093 to 95 local 173
Email: fcinches@liceo.edu.ph

Assoc. Prof. Dr. Krisda Tanchaisak

Vice President of International Affairs
Ramkhamhaeng University
Ramkhamhaeng Road, Hua Mark, Bang Kapi, Bangkok 10240. Thailand
E-mail: krisada000@yahoo.com

Dr. Sutee Sujitparapitaya

Associate Provost
Office of Institutional Analytics,
San Francisco State University,
California, USA.
Email: sutee.sujitparapitaya@gmail.com

Assoc. Prof. Dr. Koh Yit Yan

Academic Director, Program Coordinator (Mechanical Engineering)
Newcastle Australia Institute of Higher Education, Singapore
#10-02/03, Suntec Tower 4, Singapore 038986
E-mail: yityan@gmail.com

Prof. Dr. Sophia Shi-Huei Ho

Professor & Vice President for Academic Affairs,
Graduate School of Education
Ming Chuan University, Taiwan,
Taiwan
Email: shihuei@mail.mcu.edu.tw

Assoc. Prof. Dr. Leenawaty Limantara

Vice Executive Director,
The Petra Christian University Board of Trustees,
Indonesia
Email: limantara@petra.ac.id

Assoc. Prof. Dr. Vimolwan Yukongdi

Faculty Director
MBA Program,
Asian Institute of Technology, Thailand.
Email: vyukongdi@ait.ac.th

Assoc. Prof. Dr. Manika Wisessathorn

Associate Dean
Faculty of Education
Ramkhamhaeng University
2086 Ramkhamhaeng Road
Hua-mark, Bangkok, Bangkok 10240, Thailand
Telephone: +66 95 909 4173
Email: manika_w@yahoo.com

Dr. Raj Sharma

Senior Consultant in Higher Education, Melbourne, Australia.
Email: dr.rajendra.sharma@gmail.com

Prof. Dr. Reiko Yamada

Faculty of Social Studies
Director, Center for Higher Education and Student Research
Doshisha University, Kyoto, Japan
Email: ryamada@mail.doshisha.ac.jp

Asso. Prof. Dr. Moon-Heum Cho

Department of Instructional Design, Development & Evaluation (IDDE)
School of Education, Syracuse University, USA
Email: mhcho@syr.edu

Dr. Sehee Kim

Management Research Analyst Principal
Office of Evaluation and Educational Effectiveness, Arizona State University, USA
Email: Sehee.Kim@asu.edu

Prof. Dr. Alvin Sario

Professor, Philosophy and Social Sciences
University of Santo Tomas - Legazpi, Legazpi City, Philippines
Email: alvin.sario@ust-legazpi.edu.ph

Dr. Antonius Tanan

Educational Advisor, Jakarta, Indonesia,
Independent Commissioner at PT Ciputra Development & PT Metrodata Electronics Tbk E-
Email: antonius.tanan@ciputra.ac.id

Dr. Ir. Chua Yaw Long, FIMechE, MIEM, CEng

Senior Lecturer, Mechanical Engineering Dept,
College of Engineering,
Universiti Tenaga Nasional, Malaysia
E-mail: Chuayl@uniten.edu.my or mailto:yawlong76@yahoo.com

Assoc. Prof. Dr. Vo Thi Kim Anh

Acting Dean, Faculty of English
University of Danang-University of Foreign Language Studies, Vietnam
Email: vtkanh@ufl.udn.vn

Dr. Familia Novita Simanjuntak

Faculty of Teachers' Training and Education
E-mail: familia.simanjuntak@uki.ac.id

JIRSEA Key Contacts (Effective 1 June 2025)

JIRSEA (Journal of Institutional Research South East Asia) © **JIRSEA 2002**
ISSN: 1675-6061

JIRSEA Publisher Address:

Sungkyunkwan University Institute of Education
25-2 Sungkyunkwan-ro, Jongno-gu, Seoul,
Sungkyunkwan University, Republic of Korea

Contact:**Paper submission and all queries:**

Editorial Manager: jirsea@skku.edu

JIRSEA (JOURNAL OF INSTITUTIONAL RESEARCH SOUTH EAST ASIA)

JIRSEA (Journal of Institutional Research South East Asia), owned by SEAAIR (South East Asian Association for Institutional Research), is an internationally affiliated organization of AIR (Association for Institutional Research) of the U.S. and has been publishing its Journal since 2002.

JIRSEA EDITORIAL POLICIES & PROCEDURES and PUBLICATION ETHICS

(a) JIRSEA: SUBMISSION OF MANUSCRIPTS

Authors and contributors to the Journal should prepare manuscripts according to these general and COPE Publications Ethics guidelines. In June 2025, JIRSEA has also instituted two main sets of guidelines: (a) JIRSEA authorship guidelines, (b) APC (author processing fees), and (c) Declaration of Use of generative AI. These guidelines provide generally accepted academic regulations and policies for preparing the Journal for publication. Failure to follow them may result in delayed and or rejected papers.

JIRSEA considers all manuscripts in the Higher Education and Institutional Research domains on the condition that:

1. They are original, use non-discriminatory language, and are not under consideration for publication elsewhere or have been published already;
2. Upon acceptance for publication, the paper copyright transfers to the journal, allowing for open access by the academic, researcher, and general public. It is licensed under the Creative Commons Attribution 4.0 International license.

(b) JIRSEA: PUBLISHING POLICY

JIRSEA is published electronically on a biannual basis. Original research papers about institutional research and tertiary education are invited for publication.

All papers undergo two stages of review:

- a. *The preliminary review (PR)* stage to strengthen or address missing research components, and once it passes this stage, and with a new submission, it goes to;
- b. *The Double-blind (DB) reviews* as refereed by two independent persons and evaluated. The manuscript is checked for plagiarism in the PR stage using Copyleaks. The authors receive a copy of the "Similarity Report" and the Editor's comments to address the paper requirements before consideration for the DB review. The reviewers' DB review identifies the following:
 1. Significance in contributing new knowledge.
 2. Research Technical Quality with in-depth adequacy addressing key research components of Literature reviews, synthesis & analysis, Research model with identified variables/constructs and measures & methodology; statistical analysis & discussions; implications, recommendations & conclusions based on findings and references.
 3. Appropriate for the Journal with presentation clarity.

(c) JIRSEA OPEN ACCESS POLICY

JIRSEA is committed to its Open Access Practices per Creative Commons  requirements to all parties who wish to download any published papers for academic or scholastics use. All downloads must subscribe to JIRSEA's commitment to proprietary intellectual citations and referencing as per COPE Publication Ethics.

(d) JIRSEA PUBLICATION ETHICS PRACTICES AND PROTOCOLS

JIRSEA has always upheld publishing and reviewing ethics that promote and practice integrity in scholarly research and publication. **JIRSEA** subscribes to the international [COPE – Committee on Publication Ethics](#). To achieve high-level publication ethics, **JIRSEA** espouses the 10 [COPE Core Practices](#). When faced with misconduct cases, the [COPE Publication Ethics FLOWCHARTS](#) systematically guides the mitigation processes addressed. **JIRSEA** Publication Ethics also subscribes to the [Council of Scientific Editors \(2018\)](#).

(e) PUBLICATION GUIDELINES

Primary Publication Ethics guiding principles that authors/researchers, peer reviewers, and editors/editorial board must adhere to:

Authors/Researchers

- Originality and unpublished manuscript submissions in partiality or whole.
- Unethical submission of the same manuscript to more than one journal concurrently and avoiding all types of plagiarism.
- Stating conflict of interest clearly and acknowledging data sources, data availability & reproducibility, citations, and references used in the manuscript development.
- Modifications to the manuscript after submission are swiftly communicated to the Editor.

Peer Reviewers

- Reviewing and providing constructive comments to all submissions in fairness based on the intellectual content of the paper, regardless of gender, race, ethnicity, religion, citizenry, or political values of the author(s).
- Observing the policy of non-conflict or competing for interest during the review process and maintaining full confidentiality of the manuscript and the reviewer's status.
- Ensuring that the manuscript publication acceptance or rejection reasons are clearly and constructively justified and communicated to the Editor.
- [COPE Code of Practice for PEER Reviews](#) as practiced by **JIRSEA**.

Editor

- Fair & Unbiased evaluation based on the intellectual context & content of the paper regardless of gender, race, ethnicity, religion, citizenry, or political values of authors.
- Confidentiality of manuscripts and reviewers' status and identity.
- Observing non-conflict of interest and intellectual properties is a standard protocol that must be disclosed.
- Taking responsibility and accountability for publication decisions for (a) submitted manuscripts based on the reviewer's evaluation of the manuscript, (b) ethical oversights by the editorial board of ethics policies, legal restraints acting against plagiarism, libel, intellectual properties & copyright infringements, complaints & appeals, (c) peer review process and journal management, and (d) post-publication discussions & corrections.
- [COPE Codes of Conduct and Best Practices Guidelines for Journal Editors](#) as practiced by **JIRSEA**.

(f) JIRSEA: SUBMISSION OF TYPESCRIPTS

Please ensure that all submissions follow the **JIRSEA** guidelines of (a) [paper guidelines](#), (b) [forms and formatting](#), and (c) [citations and references](#). All details noted above pertaining to publication requirements are available on the Journal web link at <http://www.seaairweb.info/journal/>. To be considered for publication in JIRSEA, all papers should follow the **JIRSEA** guidelines as noted here.

(g) JIRSEA: KEY CONTACT INFORMATION

All journal submissions or queries **(Effective 1 June 2025)** about JIRSEA should be made to:

Paper submission and all queries:

Editorial Manager: jirsea@skku.edu

Other Key Contacts:

Editor: jirseaeditor@skku.edu

Associate editor: jirseareview@skku.edu

Address:

Sungkyunkwan University Institute of Education
25-2 Sungkyunkwan-ro, Jongno-gu, Seoul,
Sungkyunkwan University, Republic of Korea

Editorial

Charting Interdisciplinary Frontiers in Higher Education Research: Leadership, Digital Transformation, and Sustainable Academic Practices

This JIRSEA-UPM Special Issue Vol. 24 #1 brings together 27 accepted papers from an initial submission pool of 30, following a rigorous double-blind review process that ensures the scholarly quality and relevance of each contribution. Higher education institutions worldwide are navigating an increasingly complex landscape characterized by technological disruption, demographic shifts, internationalization pressures, and evolving stakeholder expectations. The papers assembled in this issue collectively illuminate five critical domains of inquiry: (1) educational leadership and institutional transformation, (2) digital ecosystems and technology adoption, (3) curriculum alignment and competency development, (4) student psychosocial factors and well-being, and (5) cross-cultural adaptation and inclusive pedagogies.

1. Educational Leadership and Institutional Transformation (5 articles)

Five articles examine how leadership structures, governance mechanisms, and professional development programmes shape institutional effectiveness and academic careers. Saidi, Md Rami, Tahir, and Wan Othman (Article 1) systematically review the interplay between educational leadership and students' perceptions of mobility programmes, revealing that leadership quality mediates learning environments and intercultural outcomes. Huo, Aziz, and Kunasegaran (Article 2) synthesise adoption patterns of reverse mentoring in higher education, identifying governance sponsorship, open culture, and structured routines as critical enablers for intergenerational knowledge exchange. Zulkifly, Ismail, and Alias (Article 12) evaluate a collegial leadership training programme at a Malaysian public university, demonstrating significant competency improvements and moderate-to-high transfer readiness among academic staff. Umar Baki, Mohd Pauzi, and Kunjiapu (Article 22) provide a qualitative case study of early-career academics' experiences with managerial governance and KPI systems, uncovering workload intensification and blurred work-life boundaries as persistent challenges. Saidi and Aziz (Article 29) systematically review leadership development approaches in United States higher education, proposing an integrated framework that aligns experiential learning, mentoring, and co-curricular engagement.

2. Digital Ecosystems: TPACK, Micro-Credentials, AI in Education, and Online Learning (7 articles)

As digital technologies proliferate, researchers are racing to understand their pedagogical implications and systemic integration. Abdul Karim and Mohd Puad (Article 3) conduct a global bibliometric review of two decades of TPACK research in higher education, revealing continuous growth since 2006 with notable acceleration post-2018, led by contributions from the United States, China, and Indonesia. Nor Hashim and Wong (Article 4) map the intellectual landscape of micro-credential research (2015–2025), identifying five thematic clusters and highlighting the field's structural fragmentation across pedagogical, technological, and policy domains. Ismail, Taresh, Noordin, and Jalaluddin (Article 5) systematically review AI-based

writing tools in higher education, synthesising 27 studies to identify benefits for non-native English speakers alongside concerns about academic integrity and higher-order thinking skills. Abdullah, Razali, and Fadzil (Article 9) examine factors influencing acceptance of CEFR-aligned English language curricula among 405 Malaysian undergraduates, finding that attitude and remote learning readiness significantly predict curriculum acceptance ($R^2 = 0.60$). Mohamed, Mohd Ayub, Yasin, and Che Nawi (Article 10) investigate personalised mathematics learning adoption among 120 Maldivian pre-university students, with habit, learning value, and student commitment emerging as significant unique predictors ($R^2 = 0.628$). Abd Rahman and Razali (Article 21) explore how shifting instructional modalities shape rapport and communicative competence among Malaysian ESL students, revealing qualitative differences in individual interactions, class participation, and non-verbal communication across face-to-face, hybrid, and online environments. Yu and Abdrahim (Article 28) examine Chinese international students' online learning experiences in Malaysian higher education through a Community of Inquiry and Culturally Responsive Teaching framework, finding that social presence remains the most challenging dimension, compounded by cultural norms that heighten participation hesitancy.

3. Curriculum Alignment, Competency Development, and TVET (6 articles)

Ensuring that academic programmes translate into marketable skills and sustainable competencies remains a universal imperative. Shi, Omar, and Ismail (Article 6) systematically review teacher self-efficacy in higher vocational education, identifying Asia as the primary research focus while noting limited cross-cultural exploration. Salimovna and Syed-Abdullah (Article 7) synthesise stakeholder perceptions of English for Academic Purposes curricula across 2000–2025, revealing persistent gaps in skill transferability, curriculum flexibility, and institutional support. M. Yusop, Yaakop, and Ahmad (Article 13) develop a competency-based framework for classroom assessment practices in TVET, demonstrating that technical skills and career adaptability skills significantly predict assessment practices, with Industry 4.0 generic skills exerting significant indirect influence through career adaptability mediation. Cui, Yasin, and Saharuddin (Article 14) conduct a pilot randomised controlled trial of AI-simulated ethical scenarios with 20 Chinese undergraduates, finding large effect sizes (ESSQ: $d=1.32$; SEJT: $d=1.52$) favouring AI-driven simulations over traditional ethics instruction. Muhtarova and Sulaiman (Article 19) investigate professional development needs of early-career EFL teachers in Uzbekistan through focus group discussions, identifying five priority areas translated into a needs-based PD prototype. Wang, Sulaiman, Jafri, and Wong (Article 16) examine the mediating role of cognitive reappraisal between perceived autonomy support and mental health among 220 Chinese vocational college students, finding partial mediation ($\beta = .176$, 95% CI [0.098, 0.259]) with the model explaining 48.3% of variance in mental health.

4. Student Psychosocial Factors: Self-Efficacy, Procrastination, Academic Entitlement, and Well-Being (5 articles)

Four studies illuminate the psychological underpinnings of student performance, persistence, and mental health. Siah and colleagues (implicitly referenced in the editorial structure) examine self-control and grit's effects on academic procrastination, while Samsudin, Che Amat, Amat, and Abu Bakar (Article 18) descriptively survey 1,206 first-year Malaysian public university

students, finding moderate levels of assertiveness ($M = 79.87$, $SD = 18.67$), social skills ($M = 86.49$, $SD = 24.49$), and psychological well-being ($M = 136.13$, $SD = 38.38$). Song, Che Amat, and Yasin (Article 17) investigate academic entitlement among 886 university students in Henan Province, China, revealing moderately low entitlement levels with only 33.4% scoring above three points, suggesting cultural specificity of entitlement phenomena. Samsudin, Wan Othman, Mohd Ayub, and Wan Jaafar (Article 30) examine relationships between assertiveness, social skills, and psychological well-being among the same cohort, finding both assertiveness ($\beta = .192$) and social skills ($\beta = .301$) significantly predict well-being, explaining 15.3% of variance. Mahadi, Zakaria, and Puad Mohd Kari (Article 27) conduct a narrative review of Adlerian-based strategies to support students' self-efficacy and higher education readiness, identifying encouragement, reducing inferiority feelings, promoting social interest, purposeful goal setting, and recognising personal strengths as core mechanisms.

5. Cross-Cultural Adaptation, Inclusivity, and Psychosocial Support (4 articles)

As higher education becomes increasingly internationalised, understanding cross-cultural adaptation and inclusive practices grows in importance. Alias and Azhari (Article 15) examine predictors of cross-cultural adaptability among international postgraduate students in a Malaysian research university, finding that resilience ($\beta = .468$), emotional intelligence ($\beta = .338$), and social support ($\beta = .261$) collectively explain 33% of variance in adaptability. Bahori, Dzulkfli, Fakhruddin, and Ismail (Article 8) assess knowledge, attitude, and practice regarding zakat management among 236 B40 students at UiTM Negeri Sembilan, revealing high KAP levels with most mean scores exceeding 4.5, informing institutional policies for financial support. Feng, Jamaluddin, and Mohd Puad (Article 11) examine organisational identification among 268 teachers in private higher education through a mixed-methods design, identifying moderate identification levels with belongingness highest, existence lowest, and proposing coordinated strategies for each ERG dimension. Zhang, Zulkifly, and Ismail (Article 25) qualitatively investigate gendered expectations and stereotypes among female academics in Malaysian research universities, revealing four primary manifestations: unequal emotional labour expectations, gendered task distribution, perceived authority differences, and pressure to constantly demonstrate competence, analysed through Role Congruity Theory.

6. Methodological Innovations in Educational Research (2 articles)

Two papers advance methodological rigour in higher education research. Uralova and Syed Abdullah (Article 23) document an Interview Protocol Refinement process for ELT curriculum research in Uzbekistan, demonstrating how participant and expert involvement across four stages enhances trustworthiness and authenticity of qualitative data collection. A methodological paper on qualitative instrument development provides practical guidance for designing semi-structured interview protocols.

Collectively, these 27 articles traverse the spectrum from macro-level policy and structural change to micro-level classroom practice and student cognition, all undergirded by an imperative to align higher education more closely with societal, economic, and technological demands. As universities navigate digital transformation, internationalisation pressures, and evolving student demographics, these empirical studies and systematic analyses provide both a roadmap and a

provocation: to continue interrogating how we organise, teach, and learn in an ever-more complex world. We trust that this issue sparks new collaborations, fuels policy dialogues, and inspires future research that bridges disciplines and geographies in the pursuit of educational excellence.

The key synopses of these 27 papers are as follows:

Article 1: Educational Leadership and Learning Shaping Students' Perceptions of Mobility Programmes in Higher Education: A Systematic Literature Review led by Zurina Ahmad Saidi, Ahmad Aizuddin Md Rami, Zurinah Tahir, and Wan Norhayati Wan Othman of Universiti Putra Malaysia and Universiti Kebangsaan Malaysia, this study aimed to examine challenges in learning and educational leadership alongside students' understanding and perceptions of mobility programmes. Employing a systematic literature review guided by the PRISMA framework, the authors analysed 34 peer-reviewed journal articles published between 2015 and 2025 using a narrative thematic approach. They found that students' mobility experiences are influenced by interrelated academic, cultural, institutional, and structural factors, with educational leadership playing a mediating role in shaping learning environments and student perceptions. The study contributes an integrated leadership, learning, and perception perspective to mobility research.

Article 2: Reverse Mentoring for Institutional Transformation in Higher Education: A Systematic Review of Adoption Patterns, Contextual Enablers, and Learning-Innovation Outcomes by Hongyi Huo, Faiq Aziz, and Mageswari Kunasegaran of Universiti Putra Malaysia set out to examine how reverse mentoring is adopted, implemented, and linked to organisational outcomes in higher education. Following PRISMA guidelines, searches of Scopus, Web of Science, and ERIC identified ten empirical studies. Thematic synthesis revealed four adoption patterns in higher education: governance-oriented, training-centred, curriculum-embedded, and community-partnered models. Six contextual enablers proved critical: governance sponsorship, open culture, dedicated resources, structured routines, external pressures, and evaluation mechanisms. Three outcome themes emerged: learning as foundational skill development, alignment as improved cross-level communication, and innovation as cumulative institutional change.

Article 3: Two Decades of TPACK Research in Higher Education: A Global Bibliometric Review from 2006 to 2025 by Noor Ibtisam Abdul Karim and Mohd Hazwan Mohd Puad of Universiti Putra Malaysia conducted a bibliometric analysis of TPACK-related publications in higher education from 2006 to 2025. Data retrieved from the Scopus database was analysed using descriptive bibliometric techniques and science mapping with VOSviewer. Findings indicate continuous growth in TPACK research, with notable increase after 2018 and during the post-pandemic period. The United States, China, and Indonesia emerged as leading contributors. Keyword co-occurrence analysis revealed strong connections between TPACK, higher education, digital competence, online learning, and technology integration, suggesting TPACK has evolved into a widely applied and interdisciplinary framework.

Article 4: Beyond Badges: A Bibliometric Exploration of Micro-Credential Research in Higher Education (2015–2025) by Nur Hazimah Nor Hashim and Su Luan Wong of Universiti

Putra Malaysia mapped the intellectual landscape of micro-credential research through bibliometric analysis of 571 Scopus-indexed publications. Using Bibliomagika® and VOSviewer, the authors found sustained and accelerating growth in scholarly output, particularly after 2019. The United States dominates output and citation impact, while Australia demonstrates higher citation efficiency. Thematic analysis identified five major clusters: pedagogical foundations, learner engagement, digital credentialing, instructional innovation, and alternative credential pathways. The convergence of keyword and co-authorship analyses shows the field is conceptually segmented and structurally fragmented, advancing understanding of micro-credentials as a socio-technical-policy framework.

Article 5: Use of Artificial Intelligence-Based Writing (AIW) Tools in Higher Education: Opportunities, Challenges, and Future Directions for Sustainable Academic Practices by Lilliaty Ismail, Sahar M. Taresh, Nooreen Noordin, and Ilyana Jalaluddin of Universiti Putra Malaysia systematically synthesised research published between January 2020 and August 2025 on AIW tools. Following systematic searches of Scopus, Web of Science, and Embase, 27 studies were included and analysed using a narrative thematic approach with MMAT quality assessment. Findings show AIW tools can support academic writing by providing immediate feedback, assisting non-native English speakers, and improving writing efficiency. Significant concerns relate to academic integrity, authorship attribution, equitable access, and possible negative effects on higher-order thinking skills. The review highlights the need for responsible, fair, and sustainable integration of AIW tools aligned with SDG 4.

Article 6: Self-Efficacy among Higher Vocational Education College Educators: A Systematic Literature Review by Shi Yingting, Muhd Khaizer Omar, and Normala Ismail of Universiti Putra Malaysia comprehensively reviewed literature on teacher self-efficacy in higher vocational education. The 20 included studies revealed that most research focuses on teaching contexts, particularly in specialised fields such as technical subjects and writing self-efficacy. Asia is the primary research focus, where cultural context reflects regional characteristics. Empirical research in higher vocational education remains limited, and future studies should explore cross-cultural variations to support targeted professional development and improve vocational education quality.

Article 7: Stakeholder Perceptions of EAP Curricula in Higher Education: A Systematic Literature Review by Abuova Janar Salimovna of Central Asian University and Sharifah Intan Sharina Syed-Abdullah of Universiti Putra Malaysia conducted a systematic review of stakeholder perceptions regarding English for Academic Purposes curricula. Covering empirical, peer-reviewed studies published between 2000 and 2025, searches of Scopus, ERIC, ScienceDirect, and SpringerLink followed the PRISMA framework. Findings indicate stakeholders generally recognise the value of EAP curricula in promoting academic literacy and communicative competence. Persistent gaps include limited transferability of EAP skills to discipline-specific contexts, curriculum flexibility challenges, and need for ongoing evaluation and collaboration. The review recommends participatory curriculum review processes and stronger alignment between EAP content and academic tasks.

Article 8: Assessing Knowledge, Attitude, and Practice (KAP) in Zakat Management among Higher Education B40 Students by Adibah Bahori of Universiti Teknologi MARA,

Izuli Dzulkkkfli, Fathiyah Mohd Fakhruddin, and Suriani Ismail of Universiti Putra Malaysia aimed to assess KAP among 236 B40 students regarding zakat management at UiTM Negeri Sembilan. Using a quantitative approach with Likert-scale questionnaire, descriptive and correlation analyses were conducted using SPSS. Findings indicate KAP levels are high, with most mean scores exceeding 4.5, reflecting strong awareness of zakat application procedures, positive attitudes toward effectiveness in reducing financial burdens, and consistent practices. Students expressed satisfaction with the user-friendly e-zakat system and clear communication channels, providing valuable insights for zakat institutions and policymakers.

Article 9: Factors Influencing Acceptance of CEFR-Aligned English Language Curriculum at a Malaysian University by Muhammad Fazli Abdullah of Universiti Putra Malaysia and Universiti Kebangsaan Malaysia, Fazilah Razali, and Farah Aida Fadzil of Universiti Putra Malaysia explored factors influencing university students' curriculum acceptance by examining attitude, speaking anxiety, and remote learning readiness. Using a quantitative correlational design with 405 undergraduate students, Pearson's correlation tests and multiple linear regression analyses were conducted. Findings indicated students' attitude ($r = 0.73$) and remote learning readiness ($r = 0.67$) are significant predictors of curriculum acceptance, with the regression model statistically significant and strong model fit ($R^2 = 0.60$). Insights are applicable to CEFR-alignment policies and ESL/EFL contexts globally.

Article 10: Investigating Factors Influencing Personalized Mathematics Learning Adoption among Pre-university Students at the Maldives National University: An Extended UTAUT2 Perspective by Mausooma Mohamed of Universiti Putra Malaysia and Maldives National University, Ahmad Fauzi Mohd Ayub, Maizura Yasin, and Nur Raihan Che Nawi of Universiti Putra Malaysia investigated determinants of behavioural intention to adopt Moodle-supported Personalized Mathematics Learning. A cross-sectional survey of 120 students (population frame 559) employed exploratory factor analysis ($KMO = 0.871$; 74.887% cumulative variance) and multiple regression. The model explained substantial variance in BI ($R^2 = 0.628$; adjusted $R^2 = 0.598$), with habit ($\beta = 0.484$, $p < .001$), learning value ($\beta = 0.220$, $p = .023$), and student commitment ($\beta = 0.261$, $p = .017$) emerging as significant unique predictors.

Article 11: Organizational Identification of Teachers in Private Higher Education: Insights from a Mixed-Methods Study by Xiuqiang Feng, Rahimah Jamaluddin, and Mohd Hazwan Mohd Puad of Universiti Putra Malaysia examined current status, group differences, and contextual influences on teachers' organisational identification. Based on ERG theory, a questionnaire collected quantitative data from 268 teachers, complemented by semi-structured interviews with 11 teachers. Results show teachers' organisational identification is generally moderate, with identification degree in descending order: belongingness, success, existence. Age, educational level, academic title, and establishment status may bring minor differences. The study proposes three key management steps: coordinated strategies for existence identification, supportive climate for belongingness, and resource expansion for success identification.

Article 12: From Learning to Practice: Evaluating Collegial Leadership Training and Transfer among Malaysian University Academics by Nurul Afifah Zulkifly, Ismi Arif Ismail, and Siti Noormi Alias of Universiti Putra Malaysia examined effectiveness of a collegial leadership training programme delivered to academics at a Malaysian public university. Using

data from one experimental cohort, pre- and post-training measures assessed collegial leadership competencies, with training transfer readiness captured using a validated post-training transfer scale. Repeated measures ANCOVA indicated significant improvement in collegial leadership competencies, with participants demonstrating moderate-to-high readiness levels. Regression analysis shows post-training collegial leadership significantly influences training transfer readiness, highlighting value of structured leadership training for academic staff.

Article 13: Competency-Based Framework for Classroom Assessment Practices in TVET: The Role of Technical, Industry 4.0, and Career Adaptability Skills by Siti Raudhah M.

Yusop of Ministry of Education Malaysia, Nurhidayah Yaakop of Universiti Putra Malaysia, and Nor Azlan Ahmad of Universiti Kebangsaan Malaysia investigated teacher perceptions regarding how classroom assessment supports development of students' technical skills, Industry 4.0 generic skills, and career adaptability skills. A competency-based framework was tested using PLS-SEM with responses from 418 senior TVET teachers. Findings reveal that technical skills ($\beta = 0.268$, $t = 5.039$) and career adaptability skills ($\beta = 0.389$, $t = 5.007$) have significant positive effects on classroom assessment practices. Industry 4.0 generic skills exert significant indirect influence through mediation of career adaptability skills ($\beta = 0.328$, $t = 4.934$, $p < 0.01$).

Article 14: AI-Simulated Ethical Scenarios and Ethical Sensitivity in Higher Education: A Pilot Randomized Controlled Trial by Guo Cui, Maizura Yasin, and Norzihani Saharuddin of

Universiti Putra Malaysia investigated the effect of interactive AI-simulated ethical scenarios on students' ethical sensitivity. Using a pretest-posttest randomised controlled trial design, 20 undergraduate students were randomly assigned to experimental group ($n=10$) engaging with AI-driven ethical simulations or control group ($n=10$) receiving conventional teaching. Results revealed significant improvements in experimental group's post-test scores (ESSQ: $M=4.2$, $d=1.32$; SEJT: $M=8.5$, $d=1.52$) compared to control group (ESSQ: $M=3.2$; SEJT: $M=6.2$), with large effect sizes ($p<0.001$). Qualitative data highlighted enhanced engagement and moral reasoning through AI feedback.

Article 15: Predictors of Cross-Cultural Adaptability among International Postgraduate Students in a Malaysian Research University: A Quantitative Correlational Study by Siti

Noormi Alias and Nur Aina Khadijah Azhari of Universiti Putra Malaysia examined predictive roles of emotional intelligence, resilience, and social support on cross-cultural adaptability among international postgraduate students. Grounded in Acculturation Theory, validated instruments were utilised including Trait Emotional Intelligence Questionnaire, Resilience Scale, Perceived Social Support Questionnaire, and Generic Sense of Ability to Adapt Scale. Multiple regression analysis revealed emotional intelligence, resilience, and social support collectively explained 33% of variance in cross-cultural adaptability, with resilience emerging as strongest predictor ($\beta = .468$, $p < .001$), followed by emotional intelligence ($\beta = .338$, $p < .001$) and social support ($\beta = .261$, $p < .001$).

Article 16: The Mediating Role of Cognitive Reappraisal in the Relationship between Perceived Autonomy Support and Mental Health among Vocational College Students in Shanghai, China by Bodier Wang, Tajularipin Sulaiman, Nuraini Jafri of Universiti Putra

Malaysia, and Kai Yan Wong of SEGi University examined whether cognitive reappraisal mediates the association between perceived autonomy support and mental health among 220

vocational college students in Shanghai. Grounded in Self-Determination Theory, data were collected using Autonomy Support Questionnaire, Learning Climate Questionnaire, Emotion Regulation Questionnaire, and Mental Health Continuum-Short Form. Reliability analysis indicated good to excellent internal consistency (Cronbach's $\alpha = .831-.938$). Mediation analysis indicated cognitive reappraisal partially mediated the association between perceived autonomy support and mental health ($\beta = .176$, 95% CI [0.098, 0.259]), with the model explaining 48.3% of variance in mental health.

Article 17: The Manifestation and Implications of Academic Entitlement among University Students in Henan Province, China by Guizhi Song, Muhammad Asyraf Che Amat, and Maizura Yasin of Universiti Putra Malaysia investigated the current state of academic entitlement among university students in Henan Province within the Chinese cultural context. Utilising quantitative research methods, the study sampled 886 university students. Results showed academic entitlement level was moderately low, with only 33.4% of students scoring more than three points on the academic entitlement scale. Apart from fourth-year students whose academic entitlement scores were significantly lower than other grade levels, no significant differences were observed across other demographic variables, suggesting academic entitlement may be associated with cultural context.

Article 18: A Descriptive Study of Assertiveness, Social Skills, and Psychological Well-Being among First-Year Students in Malaysian Public Universities by Harizah Izyan Samsudin, Muhammad Asyraf Che Amat of Universiti Putra Malaysia, Salleh Amat, and Abu Yazid Abu Bakar of Universiti Kebangsaan Malaysia examined levels of assertiveness, social skills, and psychological well-being among 1,206 first-year undergraduate students in Malaysian public universities. A quantitative cross-sectional descriptive survey design employed multistage sampling incorporating cluster and stratified techniques. Data were collected using Rathus Assertiveness Schedule ($\alpha = .93$), Social Skills Inventory ($\alpha = .94$), and Ryff's Psychological Well-Being Scale ($\alpha = .88$). Descriptive findings show students reported moderate levels of assertiveness ($M = 79.87$, $SD = 18.67$), social skills ($M = 86.49$, $SD = 24.49$), and psychological well-being ($M = 136.13$, $SD = 38.38$).

Article 19: Exploring the Professional Development Needs of Early-Career EFL Teachers in Uzbekistan: Implications for University-Based Teacher Training Programs by Munisa G. Muhtarova and Tajularipin Sulaiman of Universiti Putra Malaysia investigated professional development needs of early-career EFL teachers in public schools of Olmazor District, Tashkent, Uzbekistan. An exploratory-descriptive qualitative research approach was adopted, with a context-demographics questionnaire and five focus group discussions with 25 randomly selected teachers (CEFR B2+, less than five years teaching experience). Five areas of PD needs were identified: managing large mixed-ability classes, teaching under technology constraints, teachers' speaking confidence and communicative classroom practice, practical and context-relevance of PD, and follow-up support. These findings were translated into a five-module needs-based PD prototype.

Article 21: Between Face-to-Face, Hybrid and Online: How Shifting Modalities Shape Rapport and Communicative Competence among Malaysian University Students by Siti Nadhirah Abd Rahman and Abu Bakar Razali of Universiti Putra Malaysia examined how

classroom modality (face-to-face, hybrid, and online) shapes rapport development among ESL university students and how these experiences shape learners' perceived opportunities for classroom communication supporting communicative competence. An exploratory qualitative study with in-depth semi-structured interviews of 20 university students revealed shifts in individual interactions, class engagement and participation, asking and addressing questions, role of spoken communication, after-class bonding, and non-verbal communication in online and hybrid learning environments.

Article 22: Managerial Governance and Performance Culture: A Case Study of Early-Career Academics' Experiences in a Malaysian Research University by Nordahlia Umar Baki, Khairunnisa Mohd Pauzi of Universiti Putra Malaysia, and Shamuni Kunjiapu of Universiti Malaya and New Era Universiti College drew on interviews with six early-career academics (equal number of permanent and contract staff, each with less than five years of experience) to explore how they understand and respond to governance practices and KPI requirements. Grounded in Job Demands-Resources theory, three main themes were identified: structuring of daily academic tasks; perceived realism and fairness of expectations; and governance structures as enablers or constraints. Participants described KPIs as a major source of pressure intensifying workload and blurring work-life boundaries, particularly for contract staff and newly returned mothers.

Article 23: Refining Qualitative Research Instruments through Participant Involvement: An Interview Protocol Refinement Process for ELT Curriculum Research in Uzbekistan by Nozimakhon Uralova of Central Asian University and Sharifah Intan Sharina Syed Abdullah of Universiti Putra Kazakhstan and Malaysia described an Interview Protocol Refinement process as a preparatory phase for qualitative research exploring Uzbekistan university lecturers' perceptions of ELT curriculum development and evaluation methods. The paper reports on four stages: brainstorming questions aligned with research questions using matrix; external review with experts; internal review with university lecturers; and piloting interview questions. The paper demonstrates how involvement of research participants and experts enhanced rigour and quality of the research instrument.

Article 25: What It Really Feels Like: Female Academics' Perspectives of Gendered Expectations and Stereotypes in Malaysian Research Universities by Zhang Pu, Nurul Afiah Zulkifly, and Suriani Ismail of Universiti Putra Malaysia investigated the role of gender stereotypes in determining work performance and professional experiences of female academic employees. A qualitative exploratory approach centred on in-depth interviews with four female academics. The study reveals gender bias manifested in four primary ways: unequal expectations of emotional labour, gendered patterns regarding task distribution, differences in perceived authority, and pressure on women to constantly demonstrate competence. Based on Role Congruity Theory, discussion indicates these stereotype-related expectations may serve as impediments to career advancement and reduce job satisfaction.

Article 27: Adlerian-Based Strategies to Support Students' Self-Efficacy and Higher Education Readiness: A Narrative Review by Husna Adilah Mahadi of Universiti Putra Malaysia and Universiti Malaysia Terengganu, Noor Syamilah Zakaria, and Dharatun Nissa Puad Mohd Kari of Universiti Putra Malaysia synthesised existing literature on Adlerian

concepts applied to strengthen students' self-efficacy in educational settings. A total of seven empirical and conceptual studies (published between 2014 and 2025 across Iran, Indonesia, South Korea, Pakistan, and the United States) were analysed. Five core themes emerged: encouragement as fundamental mechanism; reducing inferiority feelings and fostering courage; promoting social interest and interpersonal connection; purposeful behaviour and goal setting; and recognising personal strengths to build positive self-concept.

Article 28: Learning Online through Community of Inquiry and Culturally Responsive Teaching (CoI-CRT) Framework: Insights from Chinese Students Studying in Malaysian Higher Education by Wang Yu and Nur Aira Abdrahim of Universiti Putra Malaysia examined how Chinese students in Malaysian higher education experienced online learning through the Community of Inquiry framework with integration of Culturally Responsive Teaching. Using a qualitative design, interviews were conducted with ten Chinese international students from two public universities. Findings show students navigated online learning through a personal check and balance process. Teaching presence strongly shaped experiences, with clear organisation and timely responses as key enablers. Cognitive presence was enhanced when instructors used simulations and interactive activities. Social presence emerged as the most challenging dimension, with different cultural norms heightening participation hesitancy.

Article 29: Leadership Development in United States Higher Education: A Systematic Review of Integrated Approaches for University Students by Zurina Ahmad Saidi and Mohd Faiq Abd Aziz of Universiti Putra Malaysia systematically synthesised leadership development strategies within United States higher education. Guided by the PRISMA framework, a systematic search of Scopus database (December 2025) focusing exclusively on United States contexts yielded 59 peer-reviewed journal articles. Thematic analysis identified six interconnected dimensions: experiential learning, student-led co-curricular engagement, formal leadership education, mentored leadership development, professional and soft skill development, and research and innovation-based learning. An integrated leadership development framework is proposed to inform comprehensive leadership initiatives.

Article 30: The Relationship between Assertiveness, Social Skills, and Psychological Well-Being among Public University Students in Malaysia by Harizah Izyan Samsudin, Wan Norhayati Wan Othman, Ahmad Fauzi Mohd Ayub, and Wan Marzuki Wan Jaafar of Universiti Putra Malaysia examined relationships between assertiveness, social skills, and psychological well-being, as well as predictive contributions of assertiveness and social skills to psychological well-being among 1,206 first-year undergraduate students. Using Rathus Assertiveness Schedule ($\alpha = .93$), Social Skills Inventory ($\alpha = .94$), and Ryff's Psychological Well-Being Scale ($\alpha = .88$), findings indicated assertiveness ($r = .259, p < .001$) and social skills ($r = .344, p < .001$) were significantly and positively associated with psychological well-being. Multiple regression analysis showed both assertiveness ($\beta = .192, p < .001$) and social skills ($\beta = .301, p < .001$) significantly predicted psychological well-being, explaining 15.3% of variance.

JIRSEA Co-Editor: Assoc. Prof. Teay Shawyun, Ph.D.

Table of Contents

Papers	Page
Educational Leadership and Learning Shaping Students' Perceptions of Mobility Programmes in Higher Education: A Systematic Literature Review <i>by Zurina Ahmad Saidi, Ahmad Aizuddin Md Rami, Zurinah Tahir & Wan Norhayati Wan Othman</i>	1
Reverse Mentoring for Institutional Transformation in Higher Education: A Systematic Review of Adoption Patterns, Contextual Enablers, and Learning-Innovation Outcomes <i>by Hongyi Huo, Faiq Aziz & Mageswari Kunasegaran</i>	24
Two Decades of TPACK Research in Higher Education: A Global Bibliometric Review from 2006 to 2025 <i>by Noor Ibtisam Abdul Karim & Mohd Hazwan Mohd Puad</i>	51
Beyond Badges: A Bibliometric Exploration of Micro-Credential Research in Higher Education (2015–2025) <i>by Nur Hazimah Nor Hashim & Su Luan Wong</i>	71
Use of Artificial Intelligence-Based Writing (AIW) Tools in Higher Education: Opportunities, Challenges, and Future Directions for Sustainable Academic Practices <i>by Lilliaty Ismail, Sahar M. Taresh, Nooreen Noordin & Ilyana Jalaluddin</i>	110
Self-Efficacy among Higher Vocational Education College Educators: A Systematic Literature Review <i>by Shi Yingting, Muhd Khaizer Omar & Normala Ismail</i>	144
Stakeholder Perceptions of EAP Curricula in Higher Education: A Systematic Literature Review <i>by Abuova Janar Salimovna & Sharifah Intan Sharina Syed-Abdullah</i>	166
Assessing Knowledge, Attitude, and Practice (KAP) in Zakat Management among Higher Education B40 Students <i>by Adibah Bahori, Izuli Dzulkkfli, Fathiyah Mohd Fakhruddin & Suriani Ismail</i>	190
Factors Influencing Acceptance of CEFR-Aligned English Language Curriculum at a Malaysian University <i>by Muhammad Fazli Abdullah, Fazilah Razali & Farah Aida Fadzil</i>	221
Investigating Factors Influencing Personalized Mathematics Learning Adoption among Pre-university Students at the Maldives National University: An Extended UTAUT2 Perspective <i>by Mausooma Mohamed, Ahmad Fauzi Mohd Ayub, Maizura Yasin & Nur Raihan Che Nawi</i>	248
Organizational Identification of Teachers in Private Higher Education: Insights from a Mixed-Methods Study <i>by Xiuqiang Feng, Rahimah Jamaluddin & Mohd Hazwan Mohd Puad</i>	271
From Learning to Practice: Evaluating Collegial Leadership Training and Transfer among Malaysian University Academics <i>by Nurul Afiqah Zulkifly, Ismi Arif Ismail & Siti Noormi Alias</i>	298

Competency-Based Framework for Classroom Assessment Practices in TVET: The Role of Technical, Industry 4.0, and Career Adaptability Skills <i>by Siti Raudhah M. Yusop, Nurhidayah Yaakop & Nor Azlan Ahmad</i>	328
AI-Simulated Ethical Scenarios and Ethical Sensitivity in Higher Education: A Pilot Randomized Controlled Trial <i>by Guo Cui, Maizura Yasin & Norzihani Saharuddin</i>	366
Predictors of Cross-Cultural Adaptability among International Postgraduate Students in a Malaysian Research University: A Quantitative Correlational Study <i>by Siti Noormi Alias & Nur Aina Khadijah Azhari</i>	399
The Mediating Role of Cognitive Reappraisal in the Relationship between Perceived Autonomy Support and Mental Health among Vocational College Students in Shanghai, China <i>by Bodier Wang, Tajularipin Sulaiman, Nuraini Jafri & Kai Yan Wong</i>	426
The Manifestation and Implications of Academic Entitlement among University Students in Henan Province, China <i>by Guizhi Song, Muhammad Asyraf Che Amat & Maizura Yasin</i>	455
A Descriptive Study of Assertiveness, Social Skills, and Psychological Well-Being among First-Year Students in Malaysian Public Universities <i>by Harizah Izyan Samsudin, Muhammad Asyraf Che Amat, Salleh Amat & Abu Yazid Abu Bakar</i>	480
Exploring the Professional Development Needs of Early-Career EFL Teachers in Uzbekistan: Implications for University-Based Teacher Training Programs <i>by Munisa G. Muhtarova & Tajularipin Sulaiman</i>	509
Between Face-to-Face, Hybrid and Online: How Shifting Modalities Shape Rapport and Communicative Competence among Malaysian University Students <i>by Siti Nadhirah Abd Rahman & Abu Bakar Razali</i>	549
Managerial Governance and Performance Culture: A Case Study of Early-Career Academics' Experiences in a Malaysian Research University <i>by Nordahlia Umar Baki, Khairunnisa Mohd Pauzi & Shamuni Kunjapau</i>	572
Refining Qualitative Research Instruments through Participant Involvement: An Interview Protocol Refinement Process for ELT Curriculum Research in Uzbekistan <i>by Nozimakhon Uralova & Sharifah Intan Sharina Syed Abdullah</i>	596
What It Really Feels Like: Female Academics' Perspectives of Gendered Expectations and Stereotypes in Malaysian Research Universities <i>by Zhang Pu, Nurul Afiqah Zulkifly & Suriani Ismail</i>	635
Adlerian-Based Strategies to Support Students' Self-Efficacy and Higher Education Readiness: A Narrative Review <i>by Husna Adilah Mahadi, Noor Syamilah Zakaria & Dharatun Nissa Puad Mohd Kari</i>	668
Learning Online through Community of Inquiry and Culturally Responsive Teaching (CoI-CRT) Framework: Insights from Chinese Students Studying in Malaysian Higher Education <i>by Wang Yu & Nur Aira Abdrahim</i>	693

Leadership Development in United States Higher Education: A Systematic Review of Integrated Approaches for University Students <i>by Zurina Ahmad Saidi & Mohd Faiq Abd Aziz</i>	717
The Relationship between Assertiveness, Social Skills, and Psychological Well-Being among Public University Students in Malaysia <i>by Harizah Izyan Samsudin, Wan Norhayati Wan Othman, Ahmad Fauzi Mohd Ayub & Wan Marzuki Wan Jaafar</i>	738

Educational Leadership and Learning Shaping Students' Perceptions of Mobility Programs in Higher Education: A Systematic Literature Review

ABSTRACT

Authors: Zurina Ahmad Saidi¹, Ahmad Aizuddin Md Rami^{1*}, Zurinah Tahir² and Wan Norhayati Wan Othman¹

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
²Faculty of Social Science and Humanities, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

*Corresponding author:

ahmadaizuddin@upm.edu.my

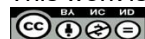
Received: 31/12/2025

First Revision: 14/1/2026

Second Revision: 15/4/2026

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Higher education mobility programs have grown steadily alongside broader internationalization initiatives, yet students' learning experiences within these programs remain shaped by a range of academic, institutional, and leadership-related challenges. Much of the existing literature addresses learning issues, educational leadership, or student perceptions separately, which limits a comprehensive understanding of how these elements interact within higher education settings. This study therefore examines challenges in learning and educational leadership alongside students' understanding and perceptions of mobility programs in higher education. Guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, a systematic literature review was conducted using the Scopus and Web of Science databases. A total of 34 peer-reviewed journal articles published between 2015 and 2025 were analyzed using a narrative thematic approach. The findings indicate that students' mobility experiences are influenced by interrelated academic, cultural, institutional, and structural factors, with educational leadership playing a mediating role in shaping learning environments and student perceptions. While mobility programs offer opportunities for intercultural learning, personal development, and emerging leadership capacity, their educational value depends strongly on leadership quality, institutional support, and inclusive learning practices. Effective educational leadership can mitigate learning challenges. The study contributes an integrated leadership, learning, perception perspective to mobility research and highlights implications for higher education leadership and mobility programs.

Keywords: educational leadership, student perceptions, learning challenges, higher education; mobility programs.

1. Introduction

Higher education has undergone significant transformation as institutions respond to increasing international connectivity, student mobility, and global interdependence. Universities across different regions have expanded mobility programs and international initiatives to enhance learning experiences, promote intercultural engagement, and prepare graduates for complex global environments. However, the expansion of these initiatives has also revealed persistent challenges related to learning processes, institutional capacity, and student experiences within diverse educational contexts (Guo and Guo, 2022; Pearson et al., 2022). These challenges suggest that internationalization in higher education is not uniformly experienced and requires closer examination.

The prior studies show that students engaged in higher education mobility often encounter difficulties that affect their academic engagement and overall learning outcomes. Differences in language use, pedagogical practices, assessment expectations, and academic cultures can complicate students' adjustment to new learning environments (Rodríguez Marconi et al., 2023; Aksezer et al., 2023). These challenges are not limited to international students alone but also affect local students and educators who participate in internationally oriented programs. As a result, learning within higher education mobility contexts involves ongoing negotiation between academic expectations and students' prior educational experiences.

In addition to learning related challenges, the literature highlights the important role of educational leadership and institutional practices in shaping student experiences. Universities vary in their ability to provide coherent support structures, inclusive policies, and effective leadership approaches that respond to increasing diversity and international engagement (Correia Monteiro et al., 2021; Monteiro et al., 2025). Recent research further emphasises that effective educational leadership in higher education is underpinned by a set of core competencies, including communication, interpersonal skills, critical thinking, emotional intelligence, and digital proficiency, which are essential in managing complex and diverse learning environments (Saidi & Abd Aziz, 2025). Leadership decisions influence how mobility programs are implemented, how students are supported, and how learning environments are structured. When institutional responses are fragmented or unclear, students may experience uncertainty and reduced engagement with their academic and social environments.

Students' understanding and perceptions of mobility programs further shape participation and outcomes in higher education. Research indicates that students often associate mobility with personal development, cultural exposure, and future career opportunities, while also expressing concerns related to cost, access, and institutional support (Guo et al., 2019; Pearson et al., 2021). These perceptions influence decision making before participation and satisfaction during and after mobility experiences. Understanding how students interpret and experience mobility programs is therefore essential for evaluating their educational value and effectiveness.

The literature on higher education mobility has generated valuable insights into students' learning experiences, institutional practices, and leadership roles. However, this body of research remains fragmented, with many studies examining learning challenges, educational leadership, or student perceptions of mobility programmes in isolation. Consequently, there is limited integrative understanding of how learning-related challenges and leadership practices together shape students' experiences and interpretations of mobility within higher education. In addition, much of the existing research is situated within specific regional, institutional, or programme contexts, which restricts broader insights into recurring patterns across diverse settings. These limitations point to the need for a systematic synthesis that brings together learning challenges, leadership and institutional factors, and student perceptions to develop a more comprehensive understanding of mobility programmes in higher education.

Addressing this gap, the present study synthesises recent scholarly evidence to examine how challenges related to learning and educational leadership interact to shape students' understanding and perceptions

of mobility programmes. By adopting an integrative perspective, this review conceptualises mobility as an educational process embedded within institutional leadership practices and learning environments.

Previous systematic reviews on student mobility have predominantly examined specific dimensions such as academic adjustment, intercultural learning, or internationalisation outcomes in isolation (King et al., 2010; de Wit & Hunter, 2015; Vande Berg et al., 2009). For instance, existing reviews have explored students' learning experiences and adaptation challenges without explicitly incorporating the role of educational or institutional leadership in shaping these processes. Similarly, other syntheses have focused on mobility outcomes, including employability and intercultural competence, yet have not sufficiently addressed how leadership practices and institutional structures influence these outcomes within higher education contexts. As a result, current systematic literature remains fragmented and offers limited explanatory power in understanding how learning challenges, educational leadership, and students' perceptions interact as an integrated system. This lack of integration constrains a more comprehensive conceptualisation of mobility as an educational process embedded within institutional practices. Therefore, this study addresses this gap by providing an integrative systematic review that examines how learning-related challenges and educational leadership jointly shape students' understanding and perceptions of mobility programmes in higher education.

1.1. Research Question and Objectives

Despite the growth of higher education mobility programmes, existing research provides limited integrated insight into how learning challenges and educational leadership jointly influence students' understanding and perceptions of mobility. While prior studies have documented academic and cultural adjustment issues and explored student experiences of mobility, fewer studies conceptualise educational leadership as a mediating structure shaping learning environments and student perceptions within mobility contexts. This lack of integration constrains the ability of higher education leaders and policymakers to design mobility programmes that are educationally coherent, inclusive, and responsive to diverse student needs.

To address this problem, this study is guided by the following research question:

How do challenges related to learning and educational leadership shape students' understanding and perceptions of mobility programmes in higher education?

In line with this question, the study pursues four objectives:

1. To synthesise learning-related challenges experienced by students participating in higher education mobility programmes.
2. To examine how educational leadership and institutional practices influence learning environments within mobility contexts.
3. To analyse patterns in students' understanding and perceptions of mobility programmes reported in the literature.
4. To develop an integrated perspective linking learning challenges, educational leadership, and student perceptions in higher education mobility.

1.2. Conceptual Framework: Educational Leadership, Learning, and Student Perceptions in Mobility Programmes

This study is underpinned by an integrative conceptual framework that positions educational leadership as a mediating structure linking learning environments and students' understanding and perceptions of mobility programmes. In higher education, educational leadership extends beyond formal managerial roles to include leadership practices that shape institutional culture, curriculum design, pedagogical approaches, and student support systems. Within mobility programs, leadership plays a critical role in determining how internationalisation is enacted in practice and how students experience learning across diverse academic and cultural environments.

Drawing on inclusive and distributed leadership perspectives, educational leadership is understood as a collective and relational process that involves academic leaders, programme coordinators, lecturers, and professional staff working collaboratively to support student learning. This perspective is also aligned with broader leadership theories in higher education, including transformational leadership, which emphasises vision, change, and the development of shared goals, as well as organisational learning perspectives that highlight institutional adaptability and continuous improvement in complex and dynamic environments. Inclusive leadership emphasises responsiveness to diversity, equity, and student voice, while distributed leadership highlights the shared responsibility for decision-making and programme implementation across institutional actors. These perspectives are particularly relevant to mobility programmes, which operate across institutional, cultural, and national boundaries and require coordinated leadership to ensure coherence and inclusivity.

From a learning perspective, mobility programmes create complex educational environments characterised by linguistic diversity, varied pedagogical traditions, and differing academic expectations. Leadership practices influence how these environments are structured, including the alignment of curricula, clarity of assessment practices, availability of academic guidance, and provision of intercultural learning opportunities. Effective educational leadership can mitigate learning challenges by fostering inclusive pedagogies, supporting staff development, and ensuring that learning environments are responsive to the needs of mobile and international students. In this study, inclusive learning practices refer to pedagogical approaches that accommodate diverse linguistic, cultural, and academic backgrounds, including differentiated instruction, accessible academic support, and student-centred engagement strategies.

Students' understanding and perceptions of mobility programmes are shaped through their interaction with these leadership-influenced learning environments. Institutional communication, programme design, and the quality of academic and social support signal to students the value and purpose of mobility experiences. When leadership practices promote transparency, inclusivity, and student-centred approaches, students are more likely to perceive mobility programmes as meaningful educational opportunities rather than as administratively complex or peripheral activities.

While the framework is presented in a linear form, it also acknowledges potential reciprocal relationships, whereby students' experiences and perceptions may influence institutional responses and leadership practices over time. Figure 1 presents the conceptual framework guiding this systematic literature review. The framework positions educational leadership as a mediating structure between broader structural and contextual conditions and students' understanding and perceptions of mobility programmes. Leadership practices shape learning environments through curriculum design, pedagogical approaches, academic support, and intercultural learning opportunities, thereby influencing how students experience and interpret mobility. Students' perceptions are thus understood not solely as individual responses, but as outcomes embedded within institutional leadership practices and learning contexts. This framework provides an integrative lens for interpreting the themes identified in the review.

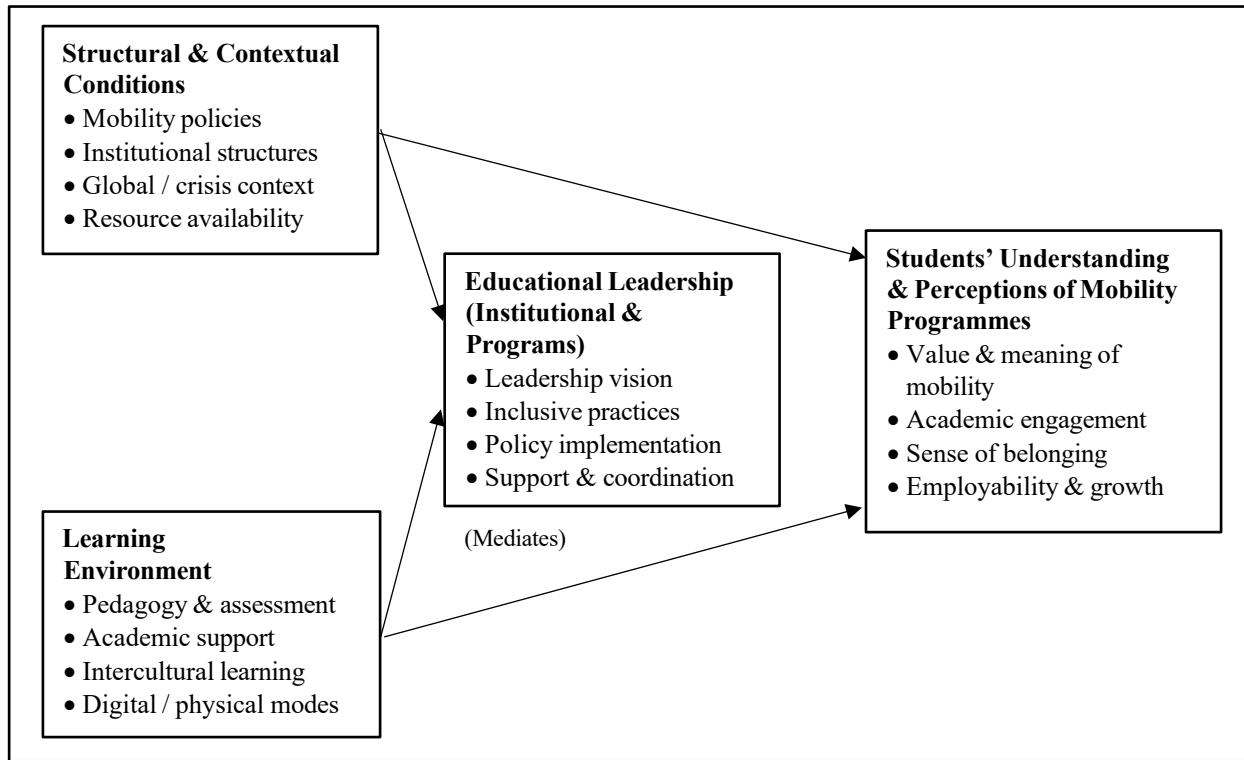


Figure 1. Conceptual Framework Illustrating the Mediating Role of Educational Leadership in Shaping Learning Environments and Students' Understanding and Perception of Mobility Programmes in Higher Education

2. Methodology

2.1. Research Design, Data Sources and Search Strategy

This study adopts a systematic literature review (SLR) guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework. The literature search was conducted between October and November 2025 using Scopus and Web of Science to identify peer-reviewed journal articles related to higher education mobility, learning challenges, educational leadership, and student perceptions. These databases were selected due to their comprehensive coverage of high-quality journals in education and the social sciences.

The search strategy employed predefined keyword combinations reflecting the core concepts of the review. Only peer-reviewed journal articles were included to ensure scholarly quality and consistency. As this study is a systematic literature review, no empirical research instruments were employed; data were derived exclusively from published journal articles relevant to the review focus.

The conceptual framework developed for this study informed both the search strategy and the analytical process. The core domains of learning challenges, educational leadership, and students' perceptions were translated into keyword combinations for database searches and subsequently operationalised as initial coding categories during thematic synthesis. This alignment ensured that the analysis remained theoretically grounded and systematically linked to the study's conceptual foundation.

The literature search was conducted using Scopus and Web of Science databases. The search was performed on 15 November 2025 using predefined keyword combinations reflecting the core concepts of mobility, higher education, programmes, and student experiences. The search strategy employed

Boolean operators (AND, OR) and field-specific queries (TITLE-ABS-KEY for Scopus and TS for Web of Science). The keywords used in the search string were as follows:

Table 1. Search String

TITLE-ABS-KEY ("student mobility" OR mobility OR "study abroad" OR exchange)
 AND TITLE-ABS-KEY ("higher education" OR university OR "tertiary education")
 AND TITLE-ABS-KEY (programme OR program OR initiative OR scheme)
 AND TITLE-ABS-KEY (learning OR perception OR experience OR outcome)
 AND PUBYEAR > 2014 AND PUBYEAR < 2026
 AND (LIMIT-TO (DOCTYPE, "ar"))
 AND (LIMIT-TO (LANGUAGE, "English"))

2.2. Inclusion and Exclusion Criteria

Articles were identified through a structured search process using Scopus and Web of Science. The search began with broad keywords related to mobility and higher education, which were subsequently combined with programme-related and experience-related terms using Boolean operators. Filters were applied to limit results to peer-reviewed journal articles in English published between 2015 and 2025 within relevant social science and education subject areas. This stepwise approach ensured a comprehensive yet focused identification of studies aligned with the review objectives.

Table 2. Core Search Concepts and Keywords

Concept	Keywords
Mobility	Mobility, exchange, study abroad, transfer
Higher education	Higher education, tertiary education, university, postsecondary
Programme	Programme, initiative, scheme, project
Experience & perception	Experience, opportunity, benefit, impact

*Database-specific search string and filters were applied within Scopus and Web of Science and are available upon request

Table 2 shows the criteria used to guide article selection and ensure methodological rigour and relevance. Only peer-reviewed journal articles were included to maintain scholarly quality, while conference papers, books, and reports were excluded due to differences in review standards. The review was limited to English-language publications to ensure consistency in analysis. Articles published between 2015 and 2025 were selected to capture contemporary research reflecting current higher education and mobility contexts. Studies were further restricted to higher education, education, and related social science fields, and only those focusing on learning, educational leadership, or student mobility were included. These criteria ensured that the final selection of articles aligned closely with the objectives of the review and supported an integrated synthesis of the literature.

Table 3. Article Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Publication type	Peer reviewed journal articles	Conference papers, books, book chapters, reviews, reports
Language	English	Non-English
Publication year	2015- 2025	Published before 2015
Subject area	Higher education, education, social sciences	Non-related fields
Study focus	Learning, leadership, or student mobility	Unrelated topics

2.3. Article Screening and Selection Process (PRISMA)

The study selection process followed the PRISMA guidelines to ensure a systematic and transparent approach to article screening and inclusion. The screening process was conducted in two stages, namely

title and abstract screening followed by full-text review. To enhance the reliability of the selection process, a second reviewer cross-checked a subset of the screened articles. Any discrepancies were discussed and resolved through consensus to ensure consistency in study selection. Figure 2 presents the PRISMA flow diagram illustrating the identification, screening, eligibility, and inclusion of studies in this systematic literature review. Records were identified from Scopus ($n = 3,993$) and Web of Science ($n = 494$), yielding a total of 4,487 records after database-level filtering. Following the removal of duplicate records ($n = 1,272$), 3,215 records were screened based on titles and abstracts.

During the eligibility stage, 620 full-text articles were assessed for relevance against the predefined inclusion and exclusion criteria. After full-text screening, 586 articles were excluded due to insufficient alignment with the review objectives. The final sample comprised 34 peer-reviewed journal articles, which formed the basis for data extraction and subsequent thematic synthesis.

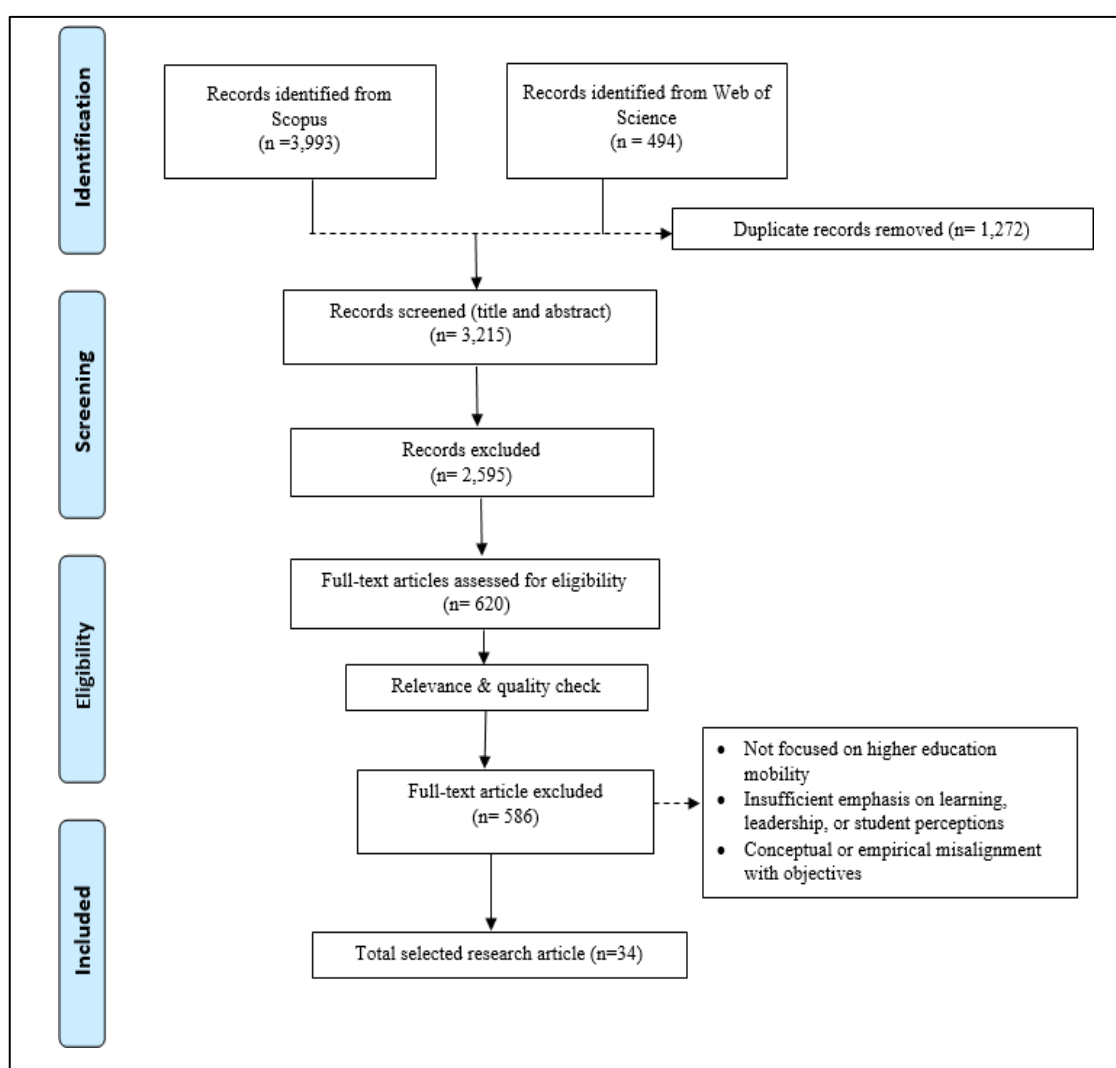


Figure 2. PRISMA Flow Diagram of the Article Selection Process

During the full-text screening stage, articles were excluded based on predefined criteria to ensure alignment with the review objectives. The main reasons for exclusion included studies not focused on higher education contexts, lack of relevance to student mobility, absence of discussion on learning or leadership dimensions, and articles that were conceptual or lacked empirical evidence. A summary of exclusion reasons and corresponding counts is presented in Table 4.

Table 4. Reasons for exclusion at full-text screening

Reason for Exclusion	Number of Articles
Not focus on higher education	150
Not related to student mobility	120
No discussion of learning or leadership	110
Conceptual papers/ lack of empirical data	90
Irrelevant outcomes (e.g., policy only, no student perspective)	116
Total	586

2.4. Quality Appraisal

To ensure the robustness and credibility of the synthesis, all included studies were subjected to a quality appraisal process. Articles were assessed based on their relevance to the review objectives, clarity of research focus and methodology, and contribution to understanding learning challenges, educational leadership, or students' perceptions within higher education mobility contexts. Only peer-reviewed journal articles meeting these criteria were retained for inclusion. This appraisal approach ensured that the final synthesis was grounded in methodologically sound and conceptually relevant literature.

2.5. Sampling Strategy, Data Extraction, and Analytical Approach

The final sample of 34 articles constituted the analytical corpus for this review. Sampling at this stage was purposive, focusing on studies that explicitly addressed learning challenges, educational leadership, or students' understanding and perceptions of mobility programmes in higher education. Data extraction focused on key study characteristics, research context, and findings relevant to the review objectives. A structured data extraction form was developed to systematically record key information from each study, including authorship, year of publication, country, research design, sample characteristics, mobility type, and key findings relevant to the review objectives.

The quality of the included studies was assessed using the Critical Appraisal Skills Programme (CASP) checklist. This tool was selected due to its suitability for evaluating qualitative, quantitative, and mixed-method studies. The appraisal criteria included clarity of research aims, appropriateness of methodology, data collection procedures, rigor of analysis, and validity of findings. Each study was systematically evaluated to ensure the inclusion of methodologically sound evidence in the synthesis. The appraisal results indicated that the majority of the included studies were of moderate to high quality, with only a small number of studies exhibiting minor methodological limitations. No studies were excluded solely based on quality appraisal, but the appraisal informed the interpretation of findings.

An inductive narrative thematic synthesis was employed to analyse the extracted data. This review adopts an interpretive epistemological stance, acknowledging that the synthesis of qualitative, quantitative, and mixed-method studies requires analytical judgement in identifying patterns and meanings across diverse forms of evidence. This approach enables the integration of heterogeneous findings while maintaining conceptual coherence in the interpretation of themes.

Initial codes were generated through repeated reading of the included studies and subsequently clustered into broader themes based on conceptual similarity and relevance. The coding process followed an iterative approach, beginning with open coding to identify initial concepts, followed by axial coding to group related categories, and finally selective coding to develop overarching themes.

Table 5. Sample Coding Framework

Code	Description	Example
Language barriers	Difficulties in academic language	Students struggled to follow lectures
Cultural adjustment	Adaptation challenges	Students faced social integration issues
Academic adaptation	Differences in pedagogy	Unfamiliar assessment methods
Institutional support	Availability of support	Mentoring improved adjustment
Leadership practices	Role of leadership	Leadership shaped programme delivery

Theme refinement involved iterative comparison across studies to identify recurring patterns and relationships. To enhance the credibility of the thematic analysis, the coding process was reviewed and refined through iterative comparison and peer discussion. This process ensured consistency in coding and interpretation of themes across the included studies. Educational leadership was analysed both as a distinct analytical dimension and as a cross-cutting influence shaping learning environments and student perceptions across themes. This approach enabled an integrated interpretation of how learning challenges and leadership practices interact within higher education mobility contexts. Greater interpretive weight was given to studies with stronger methodological rigor, as identified through the quality appraisal process, while findings from studies with limitations were interpreted with caution. The final sample of 34 studies was considered sufficient to achieve thematic saturation, as no substantially new themes emerged during the later stages of analysis.

3. Results and Discussion

Figure 3 presents a concise overview of the key challenges influencing learning, educational leadership, and mobility programs in higher education. The figure highlights academic barriers related to language, pedagogy, and learning adaptation, alongside cultural adjustment challenges involving social integration and emotional wellbeing. Leadership gaps reflect institutional and governance limitations that affect the provision of inclusive support and effective management of mobility initiatives. Mobility constraints emphasise structural barriers such as financial limitations and administrative procedures, while mobility perceptions capture how students interpret and value mobility experiences. Crisis disruptions illustrate the impact of large-scale events that interrupt learning and mobility, and developmental gains recognise the positive outcomes that can emerge despite these challenges. Collectively, these dimensions illustrate the complex and interconnected nature of challenges within higher education mobility contexts.

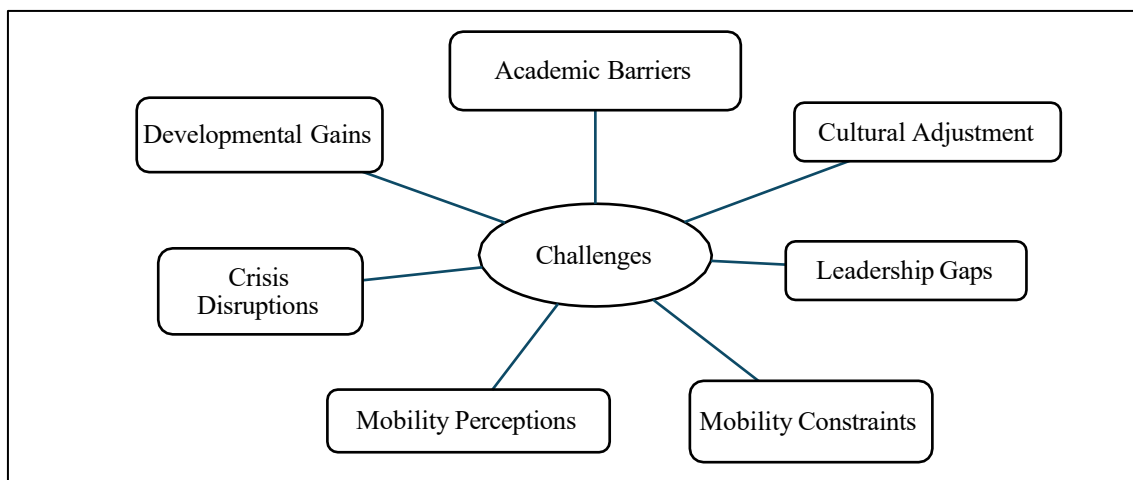


Figure 3. Key Dimensions of Challenges in Higher Education Mobility

Table 6. Key Dimensions of Challenges in Higher Education Mobility

Theme	Description	References
Academic Barriers	• Language difficulties hinder comprehension, participation, and performance • Misalignment with new pedagogical styles, curricula, and academic expectations • Limited academic guidance, research opportunities, and support structures • Difficulties adapting to online/virtual learning, including technology and time-zone issues	Rodríguez Marconi et al. (2023); Guo & Guo (2022); Aksezer et al. (2023); Monteiro et al. (2025); Pearson et al. (2022); Correia Monteiro et al. (2021); Pearson et al. (2021); Guo et al. (2019); Aksay Aksezer et al. (2023); Pearson et al. (2015); McNamara et al. (2018); Demiryontar et al. (2023); Rodríguez et al. (2020); Aksezer & Demiryontar (2021); Monteiro & Margarida (2017); Pearson & McMichael (2018); Guo (2016); Correia Monteiro (2014); Pearson et al. (2011); Aksezer et al. (2020); Guo & Guo (2019); Rodríguez Marconi et al. (2020); Monteiro et al. (2018); Pearson et al. (2017); Guo (2015); Aksezer et al. (2018); Correia Monteiro et al. (2019); Pearson et al. (2006)
Cultural Adjustment	• Students experience culture shock, marginalization, and limited social integration • Racism, discrimination, and eurocentrism affect belonging and engagement • Emotional strain from isolation, homesickness, and transitioning into unfamiliar environments	Rodríguez Marconi et al. (2023); Guo & Guo (2022); Aksezer et al. (2023); Pearson et al. (2022); Monteiro et al. (2025); Correia Monteiro et al. (2021); Pearson et al. (2021); Guo et al. (2019); Aksay Aksezer et al. (2023); Pearson et al. (2018); McNamara et al. (2017); Rodríguez et al. (2020); Aksezer & Demiryontar (2021); Monteiro & Margarida (2016); Guo (2015); Pearson et al. (2014); Correia Monteiro (2011); Aksezer et al. (2020); Pearson et al. (2019); Guo & Guo (2018); Rodríguez Marconi et al. (2020); Monteiro et al. (2017); Pearson et al. (2015); Guo (2016); Aksezer et al. (2018); Correia Monteiro et al. (2019); Pearson et al. (2006); McNamara et al. (2018); Rodríguez et al. (2014); Guo (2011); Monteiro et al. (2018)
Leadership Gaps	• Institutions lack capacity for multicultural, inclusive, and global responsive leadership • Eurocentric structures and weak support systems constrain learning and internationalization • Bureaucratic barriers, inconsistent policies, and resources limitations impede leadership effectiveness	Rodríguez Marconi et al. (2023); Aksezer et al. (2023); Pearson et al. (2022); Monteiro et al. (2025); Correia Monteiro et al. (2021); Guo et al. (2019); Pearson et al. (2021); Rodríguez et al. (2020); Aksezer & Demiryontar (2021); Guo (2018); Monteiro & Margarida (2017); Pearson et al. (2015); Correia Monteiro (2014); Aksezer et al. (2020); Guo & Guo (2019); Rodríguez Marconi et al. (2020); Pearson et al. (2018); Monteiro et al. (2016); Guo (2015); Aksezer et al. (2018); Correia Monteiro et al. (2019); Pearson et al. (2006); McNamara et al. (2017); Rodríguez et al. (2011); Guo (2016); Monteiro et al. (2018)

Mobility Constraints	- Financial constraints remain the strongest deterrent for mobility participation - Administrative, legal, and visa complexities limit access and continuity - Insufficient preparation for linguistic, academic, and intercultural demands - Limited opportunities for internships, work placements, and post-study pathways	Rodríguez Marconi et al. (2023); Guo & Guo (2022); Aksezer et al. (2023); Pearson et al. (2022); Correia Monteiro et al. (2021); Guo et al. (2019); Pearson et al. (2021); Rodríguez et al. (2020); Aksezer et al. (2021); Guo (2018); Monteiro et al. (2017); Pearson et al. (2015); Aksezer et al. (2020); Guo & Guo (2019); Rodríguez Marconi et al. (2020); Pearson et al. (2018); Monteiro et al. (2016); Guo (2015); Correia Monteiro et al. (2019); Pearson et al. (2006); Monteiro et al. (2018)
Mobility Perceptions	- Students equate internationalization primarily with physical mobility experiences - Mobility perceived as enhancing employability, cultural intelligence, and professional growth - Experiences and perceptions vary based on institutional support, inclusivity, , and destination image - Mismatch exists between official programme objectives and student expectations	Rodríguez Marconi et al. (2023); Guo & Guo (2022); Aksezer et al. (2023); Pearson et al. (2022); Correia Monteiro et al. (2021); Pearson et al. (2021); Guo et al. (2019); Rodríguez et al. (2020); Aksezer et al. (2021); Guo (2018); Monteiro et al. (2017); Pearson et al. (2015); Aksezer et al. (2020); Guo & Guo (2019); Rodríguez Marconi et al. (2020); Pearson et al. (2018); Monteiro et al. (2016); Guo (2015); Correia Monteiro et al. (2019); Pearson et al. (2006); McNamara et al. (2017); Rodríguez et al. (2011); Guo (2016); Monteiro et al. (2018)
Crisis Disruptions	- COVID-19 created major disruptions: travel bans, campus closures, online transition - Conflicts and crises damaged institutions, reduced access, and destabilized academic pathways - Reduced enrolment, emotional stress, and compromised learning quality observed	Rodríguez Marconi et al. (2023); Aksezer et al. (2023); Pearson et al. (2022); Monteiro et al. (2025); Correia Monteiro et al. (2021); Pearson et al. (2021); Rodríguez et al. (2020); Aksezer et al. (2021); Guo (2018); Monteiro et al. (2017); Pearson et al. (2015); Aksezer et al. (2020); Rodríguez Marconi et al. (2020); Pearson et al. (2018); Monteiro et al. (2016); Pearson et al. (2006)
Development Gains	- Mobility builds intercultural competence, cultural intelligence, and global mind-set - Enhances leadership, autonomy, adaptability, and reflective learning - Support career readiness and professional development despite existing challenges	Rodríguez Marconi et al. (2023); Guo & Guo (2022); Aksezer et al. (2023); Pearson et al. (2022); Correia Monteiro et al. (2021); Pearson et al. (2021); Rodríguez et al. (2020); Aksezer et al. (2021); Guo (2018); Monteiro et al. (2017); Pearson et al. (2015); Aksezer et al. (2020); Rodríguez Marconi et al. (2020); Pearson et al. (2018); Monteiro et al. (2016); Pearson et al. (2006); Monteiro et al. (2018)

3.1. Integrative Discussion

The findings of this review indicate that student mobility experiences are shaped by a complex interaction of learning challenges, institutional practices, and leadership dynamics. This integrative framework extends existing approaches by revealing how educational leadership functions as a mediating mechanism linking structural conditions and student experiences, thereby offering a more comprehensive explanation of mobility as a systemic educational process rather than isolated individual experiences. The results further demonstrate that leadership plays a critical role in shaping both learning environments and students' perceptions of mobility programmes. Viewed through the proposed conceptual framework, the findings of this review highlight educational leadership as a mediating force that shapes how learning challenges within mobility programmes are experienced and interpreted by students. Across the themes, learning-related difficulties, cultural adjustment processes, and perceptions of programme value do not emerge as isolated issues, but are embedded within leadership-influenced learning environments and institutional practices. Leadership decisions related to curriculum alignment, pedagogical support, communication, and the provision of academic and social resources consistently frame students' engagement with mobility and inform how such experiences are understood as either enriching or disruptive. This integrative perspective underscores that students' perceptions of mobility programmes are not solely individual responses to unfamiliar contexts, but are outcomes shaped by leadership practices that structure learning conditions, signal institutional priorities, and mediate the relationship between mobility participation and educational outcomes. This broader interpretation is further supported by research demonstrating that experiential engagement through structured programmes enhances students' readiness, active participation, and developmental outcomes, thereby reinforcing the role of practice-based learning within higher education contexts (Saidi & Abd Aziz, 2025; Saidi et al., 2025).

3.2. Theme 1: Academic Barriers

Academic and pedagogical challenges are among the most frequently discussed issues in the literature on higher education mobility and international learning. Across the prior studies, learning difficulties are not limited to subject knowledge alone but are closely linked to language use, teaching approaches, assessment practices, and students' ability to adapt to unfamiliar academic cultures. These challenges shape how students engage with learning environments and influence their academic confidence, participation, and overall educational experience.

A recurring concern highlighted in the literature is language-related difficulty, particularly for students studying in a second or third language. Several studies report that limited academic language proficiency affects students' ability to follow lectures, participate in discussions, complete written assignments, and engage critically with course content (Guo & Guo, 2022; Pearson et al., 2022; Correia Monteiro et al., 2021). Even when students possess conversational language skills, academic writing and discipline-specific terminology remain significant obstacles. These language challenges often result in reduced classroom interaction and increased reliance on surface learning strategies rather than deep engagement.

Another major academic challenge relates to differences in pedagogical styles and learning expectations. Students frequently struggle to adapt to teaching approaches that emphasise independent learning, critical thinking, and self-directed study, particularly when coming from more structured or teacher-centred educational systems (Rodríguez Marconi et al., 2023; Aksezer et al., 2023; Pearson et al., 2021). Assessment methods such as continuous evaluation, group work, and reflective assignments are often unfamiliar and can create uncertainty about academic standards and performance expectations.

The prior studies also highlight limited academic guidance and support as a persistent issue. Many students report uncertainty about academic norms, feedback processes, and research expectations, particularly at postgraduate level (Monteiro et al., 2025; Guo et al., 2019). In some cases, insufficient supervision, unclear communication with lecturers, or lack of structured academic mentoring

contributes to confusion and reduced learning confidence. These challenges are amplified when institutional support mechanisms are not clearly communicated or easily accessible.

With the expansion of online and blended learning, especially during periods of disruption, digital and virtual learning challenges have become more visible. Students face difficulties related to time zone differences, reduced interaction, technological limitations, and a lack of immediate feedback in virtual classrooms (Pearson et al., 2022; Aksezer et al., 2023). While online learning offers flexibility, several studies indicate that it can also intensify feelings of academic isolation and hinder collaborative learning, particularly for students unfamiliar with digital learning platforms.

The literature further suggests that academic challenges are not experienced equally by all students. Background factors such as prior educational experience, cultural capital, and socioeconomic position influence how well students adapt to new learning environments (Correia Monteiro et al., 2021; Rodríguez Marconi et al., 2023). Students from underrepresented or non-traditional backgrounds often face greater difficulty navigating academic systems that assume prior familiarity with dominant educational norms.

In summary, the prior studies demonstrate that academic and pedagogical challenges are complex and interconnected. Language barriers, unfamiliar teaching methods, limited academic support, and the demands of digital learning collectively shape students' learning experiences. Addressing these challenges requires not only individual student effort but also institutional awareness and pedagogical responsiveness to diverse learning needs across higher education contexts.

3.3. Theme 2: Cultural Adjustment

Social, cultural, and emotional adjustment challenges emerge as one of the most prominent themes in the literature on higher education mobility and international learning. Across the prior studies, students' academic experiences are closely intertwined with their ability to adapt socially and culturally to new environments. These challenges extend beyond the classroom and shape students' sense of belonging, emotional well-being, and overall engagement with higher education institutions.

A widely reported issue is difficulty in social integration, particularly among international and mobile students. Many studies describe feelings of isolation, loneliness, and exclusion, especially during the early stages of transition (Guo & Guo, 2022; Pearson et al., 2022; Correia Monteiro et al., 2021). Limited interaction with local students, weak peer networks, and unfamiliar social norms often prevent meaningful social connections from forming. As a result, students may remain within small cultural or linguistic groups, which restricts broader intercultural engagement.

Closely linked to social adjustment are cultural differences and misunderstandings. Several studies highlight how unfamiliar academic cultures, communication styles, and social expectations can lead to discomfort and uncertainty (Rodríguez Marconi et al., 2023; Aksezer et al., 2023; Pearson et al., 2021). Students often struggle to interpret classroom behaviours, staff–student relationships, and informal interaction norms, which can affect their confidence and willingness to participate. These challenges are intensified when institutions do not explicitly address cultural orientation or intercultural learning.

Experiences of discrimination, racism, and Eurocentrism are also frequently reported in the literature. Some students describe feeling marginalised due to language accent, nationality, race, or cultural background, which negatively affects their emotional well-being and sense of safety (Correia Monteiro et al., 2021; Guo et al., 2019). Eurocentric curricula and dominant cultural practices can further reinforce feelings of exclusion, particularly for students from non-Western contexts. Such experiences may discourage participation and weaken students' identification with the institution.

Emotional challenges such as homesickness, anxiety, and stress are common across mobility experiences. Separation from family and familiar support systems, combined with academic pressure and cultural adaptation, creates emotional strain for many students (Pearson et al., 2022; Monteiro et

al., 2025). Several studies note that these emotional difficulties are often hidden and underreported, as students may be reluctant to seek help or may be unaware of available support services.

The literature also shows that crisis contexts, particularly during periods of disruption, amplify social and emotional challenges. Reduced face-to-face interaction, limited campus activities, and reliance on online communication increase feelings of isolation and emotional fatigue (Aksezer et al., 2023; Pearson et al., 2021). In such situations, students report a weakened sense of community and reduced opportunities for informal social learning, which are central to intercultural experiences.

Overall, the prior studies demonstrate that social, cultural, and emotional adjustment challenges play a critical role in shaping students' higher education experiences. Difficulties in social integration, cultural understanding, and emotional well-being can undermine both academic engagement and personal development. Addressing these challenges requires intentional institutional efforts to foster inclusive environments, strengthen peer interaction, and provide accessible emotional and intercultural support for diverse student populations.

3.4. Theme 3: Leadership Gaps

Institutional and educational leadership challenges form a central strand in the literature on higher education mobility and internationalisation. The prior studies consistently show that students' learning and mobility experiences are strongly shaped by how institutions are organised, led, and governed. Rather than individual shortcomings, many difficulties faced by students are linked to institutional structures, leadership priorities, and the capacity of higher education systems to respond to increasing diversity and global engagement.

One major challenge identified across the literature is the limited institutional readiness to support diverse student populations. Several studies point out that universities often promote internationalization rhetorically, yet lack coherent strategies to support multicultural learning environments in practice (Rodríguez Marconi et al., 2023; Pearson et al., 2022; Correia Monteiro et al., 2021). This gap is reflected in inconsistent support services, fragmented policies, and unclear communication channels, which leave students uncertain about where and how to seek academic or personal assistance.

Closely related to this issue is the persistence of Eurocentric institutional cultures and curricula. Many studies highlight how dominant educational norms continue to privilege Western perspectives, teaching methods, and knowledge systems, limiting meaningful inclusion of diverse epistemologies (Guo et al., 2019; Aksezer et al., 2023). Educational leaders face challenges in reconceptualizing curricula and pedagogical approaches to reflect multicultural realities, particularly within institutions shaped by long-standing traditions and regulatory constraints.

The literature also draws attention to bureaucratic and administrative barriers within higher education institutions. Complex procedures, rigid regulations, and slow decision-making processes often hinder effective implementation of internationalization and mobility initiatives (Monteiro et al., 2025; Pearson et al., 2021). From students' perspectives, administrative inflexibility can translate into delays, confusion, and frustration, particularly in areas such as course registration, credit recognition, and academic progression.

Leadership capacity and vision emerge as another critical concern. Several studies emphasize that educational leaders are increasingly required to manage multilingual, multicultural, and transnational contexts, yet may lack adequate training or institutional support to do so effectively (Rodríguez Marconi et al., 2023; Aksezer et al., 2023). Leadership approaches that prioritise compliance or rankings over student experience may further weaken the ability of institutions to respond to learners' diverse needs.

Resource constraints represent an additional institutional challenge highlighted in the prior studies. Limited funding, staffing shortages, and uneven allocation of resources restrict the development of

inclusive programmes, staff training, and student support mechanisms (Guo et al., 2019; Correia Monteiro et al., 2021). These constraints are particularly evident in institutions operating within unequal global partnerships, where power imbalances can shape priorities and limit local agency in decision-making.

In summary, the literature indicates that institutional and leadership challenges play a decisive role in shaping higher education experiences. Weak support structures, Eurocentric practices, bureaucratic complexity, and limited leadership capacity collectively constrain effective internationalisation and student engagement. Addressing these challenges requires strategic leadership that prioritises inclusivity, flexibility, and responsiveness to the realities of an increasingly diverse higher education landscape.

3.5. Theme 4: Mobility Constraints

Structural and systemic barriers are widely identified as key constraints shaping participation in mobility programs in higher education. Across the prior studies, these barriers operate beyond individual motivation and are embedded within financial systems, administrative arrangements, legal frameworks, and institutional policies. As a result, access to mobility opportunities is uneven, and students' experiences are often shaped by constraints that lie outside their direct control.

One of the most consistently reported barriers is financial constraint. Many studies highlight the high costs associated with mobility, including tuition fees, living expenses, travel costs, and insurance, which discourage participation or limit the duration of mobility experiences (Guo & Guo, 2022; Pearson et al., 2022; Correia Monteiro et al., 2021). Even when scholarships or funding schemes are available, they are often insufficient or highly competitive, creating unequal access for students from lower socioeconomic backgrounds.

Alongside financial issues, administrative and legal complexities emerge as major systemic challenges. Students frequently encounter difficulties related to visa procedures, documentation requirements, credit recognition, and institutional regulations (Rodríguez Marconi et al., 2023; Pearson et al., 2021). These processes are often described as time-consuming, unclear, and inconsistently applied across institutions. Such bureaucratic barriers can delay mobility participation, create stress, and in some cases discourage students from engaging in mobility programs altogether.

The literature also points to limited opportunities for work, internships, and post-study pathways as a significant structural constraint. Several studies note that restrictions on employment during or after studies reduce the perceived value of mobility programs, particularly for students who prioritise employability and career development (Guo et al., 2019; Monteiro et al., 2025). When mobility experiences are not clearly linked to professional opportunities, students may question their long-term benefits despite personal or cultural gains.

Another systemic issue concerns unequal institutional capacity and global power imbalances. Some studies highlight how partnerships between institutions are shaped by unequal resources, regulatory frameworks, and decision-making power, often privileging institutions from more developed contexts (Correia Monteiro et al., 2021; Rodríguez Marconi et al., 2023). These asymmetries can limit reciprocity in mobility programs and restrict opportunities for students from less-resourced institutions.

Preparation and access to information also emerge as structural barriers. Several studies report that students receive insufficient guidance before departure, particularly regarding academic expectations, legal procedures, and cultural adaptation (Pearson et al., 2022; Aksezer et al., 2023). Inadequate preparation increases uncertainty and places additional pressure on students to navigate complex systems independently, often relying on informal networks rather than institutional support.

In summary, the prior studies show that structural and systemic barriers significantly shape who can participate in mobility programs and how those experiences unfold. Financial limitations, bureaucratic

complexity, restricted employment opportunities, and institutional inequalities collectively constrain access and outcomes. Addressing these barriers requires coordinated policy action, clearer administrative processes, and institutional commitment to making mobility programs more inclusive and accessible across higher education systems.

3.6. Theme 5: Mobility Perceptions

Understanding and perceptions of mobility programs play a crucial role in shaping students' decisions to participate and their overall experiences in higher education. Across the prior studies, mobility is commonly understood not merely as an academic activity but as a personal, cultural, and professional journey. Students' perceptions are shaped by expectations, lived experiences, and the level of institutional support they receive before, during, and after mobility.

A dominant perception identified in the literature is the strong association between internationalisation and physical mobility. Many students equate internationalisation primarily with studying abroad or participating in exchange programmes, often overlooking virtual or non-traditional forms of mobility (Guo & Guo, 2022; Pearson et al., 2022). This narrow understanding influences how students evaluate the value of mobility programs and can limit interest in alternative international learning opportunities.

Mobility programs are frequently perceived as pathways to personal growth and cultural exposure. Several studies report that students value opportunities to experience different cultures, develop intercultural awareness, and broaden their worldviews (Rodríguez Marconi et al., 2023; Aksezer et al., 2023; Correia Monteiro et al., 2021). These experiences are often described as transformative, contributing to greater self-confidence, independence, and openness to diversity.

At the same time, perceptions of mobility are closely linked to employability and career development. Many students view participation in mobility programs as a way to enhance their future job prospects, acquire transferable skills, and gain international experience valued by employers (Guo et al., 2019; Monteiro et al., 2025). When mobility is perceived to align with professional goals, students tend to report higher levels of satisfaction and motivation.

However, the literature also highlights mismatches between expectations and realities. Some students enter mobility programs with idealised expectations, only to encounter academic, cultural, or administrative challenges that alter their perceptions (Pearson et al., 2021; Rodríguez Marconi et al., 2023). When institutional support is weak or communication is unclear, these mismatches can lead to disappointment and reduced perceived value of the mobility experience.

Institutional context plays a significant role in shaping perceptions. Studies emphasise that positive perceptions are reinforced when institutions provide clear guidance, inclusive environments, and meaningful academic integration (Correia Monteiro et al., 2021; Aksezer et al., 2023). Conversely, limited support, weak recognition of mobility outcomes, or lack of connection to academic programmes can diminish students' appreciation of mobility initiatives.

In summary, the prior studies demonstrate that students' understanding and perceptions of mobility programs are multifaceted and shaped by personal aspirations, cultural experiences, and institutional practices. While mobility is widely viewed as beneficial for personal and professional development, its perceived value depends on realistic expectations, effective support structures, and clear alignment with academic and career pathways within higher education.

3.7. Theme 6: Crisis Disruptions

Global crises have had a profound impact on learning and mobility in higher education, reshaping how students engage with academic environments and international opportunities. The prior studies show that crises, particularly large-scale disruptions, expose vulnerabilities within higher education systems and significantly affect students' academic progression, mobility participation, and overall well-being.

These impacts extend beyond temporary disruption and have lasting implications for how learning and mobility are organised. One of the most visible effects of global crises is the disruption of physical mobility. Travel restrictions, border closures, and institutional shutdowns have directly limited students' ability to participate in exchange programmes and study abroad initiatives (Pearson et al., 2022; Rodríguez Marconi et al., 2023). Several studies note a sharp decline in mobility participation, with students either postponing or cancelling planned international experiences. For many, these disruptions altered academic pathways and delayed graduation timelines.

The transition to online and remote learning emerges as a central response to crisis conditions. While digital platforms enabled continuity of instruction, many students experienced difficulties adapting to fully online environments (Aksezer et al., 2023; Pearson et al., 2021). Challenges included reduced interaction with lecturers and peers, limited opportunities for collaborative learning, and increased cognitive and emotional fatigue. The absence of face-to-face engagement also weakened the intercultural dimension of mobility experiences, which is often considered a core benefit of international education. Emotional and psychological challenges are strongly emphasised across the studies. Students reported heightened levels of stress, anxiety, and uncertainty, particularly when crises disrupted both academic routines and personal lives (Monteiro et al., 2025; Correia Monteiro et al., 2021). Feelings of isolation were intensified by campus closures and social distancing measures, especially for students who were separated from family or support networks. These emotional pressures often affected concentration, motivation, and academic performance.

The prior studies also highlight institutional challenges in crisis response. Several studies suggest that universities were unevenly prepared to manage large-scale disruptions, leading to inconsistent communication, unclear policies, and unequal access to support services (Rodríguez Marconi et al., 2023; Pearson et al., 2021). Students' experiences during crises were shaped not only by the severity of the disruption but also by how effectively institutions responded to emerging needs. Despite these challenges, some studies identify adaptive practices and learning opportunities arising from crisis conditions. Hybrid models combining online and physical mobility, increased flexibility in assessment, and expanded use of virtual collaboration tools were viewed positively by some students (Aksezer et al., 2023; Pearson et al., 2022). However, these adaptations were generally seen as complementary rather than substitutes for in-person mobility and interaction.

In summary, the prior studies demonstrate that global crises significantly disrupt learning and mobility in higher education. Restrictions on movement, rapid shifts to online learning, emotional strain, and uneven institutional responses collectively shape students' experiences. While adaptive strategies have mitigated some impacts, the findings suggest a need for more resilient, flexible, and student-centred approaches to managing future crises within higher education systems.

3.8. Theme 7: Development Gains

Opportunities for leadership development emerge as an important, though less uniformly discussed, outcome of mobility experiences in higher education. Across the prior studies reviewed, mobility is not only associated with academic or cultural exposure but also with the development of personal and professional capacities that are closely linked to leadership. These opportunities are shaped by students' engagement with unfamiliar environments, responsibility-taking, and reflective learning during mobility experiences.

A recurring theme in the literature is the development of intercultural competence and global awareness as foundational leadership attributes. Several studies note that exposure to diverse cultural contexts encourages students to develop sensitivity, adaptability, and cross-cultural communication skills (Guo & Guo, 2022; Pearson et al., 2022). Through interaction with peers from different backgrounds, students learn to navigate differences, manage misunderstandings, and engage constructively in multicultural settings, all of which are essential for leadership in globalised contexts.

Mobility experiences are also linked to increased autonomy and self-efficacy. Students frequently report gaining confidence through managing daily challenges such as independent living, academic decision-making, and problem-solving in unfamiliar systems (Rodríguez Marconi et al., 2023; Aksezer et al., 2023). These experiences contribute to a stronger sense of responsibility and initiative, qualities commonly associated with emerging leadership capacity. Several studies highlight the role of mobility in fostering reflective and critical thinking. Encountering different educational practices and social norms encourages students to question assumptions and reflect on their own values and behaviours (Correia Monteiro et al., 2021; Pearson et al., 2021). This reflective process supports personal growth and enhances students' ability to lead with empathy and awareness, particularly in culturally diverse environments.

Leadership development is also connected to career orientation and professional identity formation. Some studies suggest that mobility helps students clarify career goals, develop transferable skills, and strengthen professional networks (Monteiro et al., 2025; Guo et al., 2019). When mobility experiences are aligned with academic programmes or professional pathways, students are more likely to recognise their leadership potential and apply acquired skills in future roles. However, the literature also indicates that leadership development through mobility is not automatic. Opportunities vary depending on programme design, duration, institutional support, and students' level of engagement (Pearson et al., 2022; Aksezer et al., 2023). Without structured reflection, mentoring, or intentional leadership-oriented activities, the developmental benefits of mobility may remain implicit or unevenly realised.

In summary, the prior studies show that mobility programs offer meaningful opportunities for leadership development through intercultural learning, autonomy, reflection, and professional growth. While not all students experience these outcomes to the same extent, mobility can serve as a powerful platform for nurturing future leaders when supported by intentional programme design and institutional commitment within higher education.

4. Implications

4.1. Implications for Educational Leadership

The findings of this review suggest that educational leadership plays a critical role in shaping the quality and educational value of higher education mobility programmes. Leaders at institutional and programme levels influence how learning challenges are anticipated, addressed, and communicated to students through decisions related to curriculum alignment, pedagogical support, and academic coordination. Effective leadership is therefore essential in ensuring that mobility programmes are not positioned solely as internationalisation initiatives, but are embedded within coherent learning structures that support diverse student needs. This highlights the importance of leadership practices that are inclusive, responsive, and attentive to the academic and developmental dimensions of student mobility.

4.2. Implications for Institutional Practice

At the level of institutional practice, the review underscores the need for mobility programmes to be supported by structured academic guidance, clear communication of learning expectations, and accessible support mechanisms for mobile students. Institutions should ensure that teaching staff and programme coordinators are equipped to work with students from diverse academic and cultural backgrounds, and that assessment practices are transparent and aligned across contexts. Institutions should implement structured academic support mechanisms such as pre-departure academic orientation, early-stage mentoring programmes, and continuous academic advising throughout the mobility period. Strengthening coordination between academic units and mobility offices may enhance students' learning experiences and reduce uncertainty during mobility, thereby contributing to more positive student perceptions of programme value.

4.3. Policy Implication

The findings of this study also carry important implications for higher education policy. At the national and institutional policy levels, there is a need to move beyond viewing student mobility as a purely administrative or internationalisation agenda, towards recognising it as a structured educational process that requires coherent leadership and academic integration. Policymakers, including ministries of higher education and accreditation bodies, should develop comprehensive frameworks that ensure the quality and inclusivity of mobility programmes. This includes establishing clear guidelines for academic credit transfer, standardising support structures for incoming and outgoing students, and embedding student wellbeing considerations within mobility policies.

In addition, funding agencies should prioritise equitable access to mobility opportunities by addressing financial barriers and supporting underrepresented student groups. Policies that encourage collaboration between institutions, strengthen leadership capacity, and promote inclusive practices are essential to ensure that mobility programmes deliver meaningful educational outcomes rather than merely symbolic internationalisation goals.

4.4. Implications for Future Research

This study is subject to several limitations that should be considered when interpreting the findings. First, the review was limited to articles indexed in Scopus and Web of Science, which may have resulted in the underrepresentation of relevant studies published in other databases or regional journals. This may bias the findings towards perspectives from more established research contexts. Second, the restriction to English-language publications may have excluded important contributions from non-English-speaking regions, particularly in areas with high levels of student mobility such as Asia and parts of Europe. As a result, the findings may reflect a disproportionate emphasis on Anglophone perspectives and institutional contexts.

Third, the use of an interpretive narrative synthesis approach introduces a degree of subjectivity in the coding and interpretation of themes, despite efforts to ensure analytical rigor through systematic procedures. While this approach enables the integration of diverse forms of evidence, it may also influence how findings are synthesised and interpreted. These limitations suggest that the findings should be interpreted with caution, particularly when generalising across different institutional and cultural contexts. Future research may benefit from including a broader range of databases, multilingual sources, and complementary methodological approaches to enhance the comprehensiveness and robustness of the evidence base.

5. Limitations

This systematic literature review is subject to several limitations. The analysis was limited to studies indexed in Scopus and Web of Science, which, although comprehensive, may not capture all relevant literature from other databases or regional sources. In addition, only English-language journal articles were included, which may restrict the representation of perspectives from non-English-speaking contexts. The review also employed a narrative thematic synthesis approach, which is appropriate for integrating diverse studies but relies on interpretive judgement and does not seek to establish causal relationships or quantify effects. These limitations notwithstanding, the review provides a coherent and conceptually grounded synthesis of how educational leadership shapes learning environments and students' understanding and perceptions of higher education mobility programmes.

6. Conclusion

This systematic literature reviews synthesised evidence on learning challenges, educational leadership, and students' understanding and perceptions of mobility programmes in higher education. The findings show that academic barriers, cultural adjustment processes, institutional support structures, and

students' interpretations of mobility experiences are closely interconnected. Across these themes, educational leadership emerges as a central factor shaping how learning challenges are experienced and how mobility programmes are understood, highlighting that mobility outcomes are not determined solely by individual student characteristics or programme design.

Through an integrative leadership perspective, the review demonstrates that leadership actions at institutional and programme levels mediate learning environments and influence students' perceptions of mobility. Decisions related to curriculum alignment, pedagogical support, assessment clarity, and coordination of academic and support services play a critical role in determining whether mobility is experienced as educationally coherent and developmentally meaningful. By positioning educational leadership as a mediating force across learning challenges and student perceptions, this study contributes a conceptual understanding of mobility as an educational process embedded within leadership and institutional practices, offering a foundation for strengthening the quality and impact of higher education mobility programmes.

7. Ethic Statement

This study is a systematic literature review and did not involve human participants, primary data collection, or identifiable personal data. As such, formal ethics approval was not required.

8. Data Availability Statement

The data analysed in this study were derived from publicly available peer-reviewed journal articles indexed in Scopus and Web of Science. No new datasets were generated or analysed.

9. Generative AI Statement

During the preparation of this manuscript, AI-based tools were used to support specific aspects of the writing and analytical process. ChatGPT 4.0 assisted in refining language clarity and improving overall readability, while Elicit was used to help organise and interpret thematic patterns emerging from the reviewed studies. All outputs generated through these tools were carefully reviewed, critically evaluated, and substantively revised by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the published work.

References

- Abd Rahim, N. A. (2021). Adjustment experiences of international graduate students to life in the United States. *Pertanika Journal of Social Sciences & Humanities*, 29(4), 2243–2262. <https://doi.org/10.47836/pjssh.29.4.08>
- Abrantes, P., Silva, A. P., Backstrom, B., Neves, C., Falé, I., Jacquinet, M., Ramos, M. R., Magano, O., & Henriques, S. (2022). Transversal competences and employability: The impacts of distance learning university according to graduates' follow-up. *Education Sciences*, 12(2), 65. <https://doi.org/10.3390/educsci12020065>
- Aksay Aksezer, E., Demiryontar, B., Dorrity, C., & Mescoli, E. (2023). International student experiences in three superdiverse higher education institutions: Institutional policies and intersectionalities. *Social Sciences*, 12(10), 544.
- Ayoubi, R. M., Solwa, B., Ayoubi, M. T., & Crawford, M. (2025). Bridging the gap: Academic performance and experiences of technical college students in top-up university degree programme. *Cogent Education*, 12(1), 2441059. <https://doi.org/10.1080/2331186X.2024.2441059>
- Bañegil-Palacios, T. M., & Sánchez-Hernández, M. I. (2018). The challenge to foster foreign students' experiences for sustainable higher educational institutions. *Sustainability*, 10(2), 495. <https://doi.org/10.3390/su10020495>
- Blankvoort, N., Kaelin, V. C., Poerbodipoero, S., & Guidetti, S. (2019). Higher education students' experiences of a short-term international programme: Exploring cultural competency and professional development. *Educational Research. Advance online publication*. <https://doi.org/10.1080/00131881.2019.1632725>
- Brutt-Griffler, J., Nurunnabi, M., & Kim, S. (2020). International Saudi Arabia students' level of preparedness: Identifying factors and maximizing study abroad experience using a mixed-methods approach. *Journal of International Students*, 10(4), 976–1004. <https://doi.org/10.32674/jis.v10i4.839>
- Curtis, T., & Ledgerwood, J. R. (2018). Students' motivations, perceived benefits and constraints towards study abroad and other international education opportunities. *Journal of International Education in Business*, 11(1), 63–78. <https://doi.org/10.1108/JIEB-01-2017-0002>
- De Wit, H., & Hunter, F. (2015). The future of internationalization of higher education in Europe. *International Higher Education*, (83), 2-3. <https://doi.org/10.6017/ihe.2015.83.9073>
- Dinh, N. B. K., Zhu, C., Caliskan, A., & Cheng, Z. (2023). Academic leadership development: An exploratory study on the impacts at the institutional level based on an Erasmus+ capacity building project. *SAGE Open*, 13(4), 1–19.
- Eijssvogels, T. M. H. E., Goorden, R., van den Bosch, W., & Hopman, M. T. E. (2015). The binding study advice in medical education: A 2-year experience. *Perspectives on Medical Education*, 4(1), 39–42. <https://doi.org/10.1007/s40037-015-0164-1>
- García, S., Harker, A., & Cuartas, J. (2019). Building dreams: The short-term impacts of a conditional cash transfer program on aspirations for higher education. *International Journal of Educational Development*, 64, 48–57. <https://doi.org/10.1016/j.ijedudev.2018.12.006>
- Guo, Y., & Guo, S. (2022). Internationalization of Canadian teacher education: Teacher candidates' experiences and perspectives. *ECNU Review of Education*, 5(3), 425–449.

- Han, Y., Li, W., Bao, M., & Cao, X. (2020). An investigation of the experiences of working with multilingual international students among local students and faculty members in Chinese universities. *Sustainability*, 12(16), 6419. <https://doi.org/10.3390/su12166419>
- Henke, J. (2019). Third mission as an opportunity for professionalization in science management. *Publications*, 7(4), 62. <https://doi.org/10.3390/publications7040062>
- Iskhakova, M., Bradly, A., Whiting, B., & Lu, V. N. (2022). Cultural intelligence development during short-term study abroad programmes: The role of cultural distance and prior international experience. *Studies in Higher Education*, 47(8), 1694-1711. <https://doi.org/10.1080/03075079.2021.1957811>
- Khalid, M. S., Zhanyong, Q., & Bibi, J. (2022). The impact of learning in a diversified environment: Social and cognitive development of international students for global mind-set. *European Journal of Training and Development*, 46(5-6), 373-389. <https://doi.org/10.1108/EJTD-12-2020-0175>
- Khoo, S., Mucha, W., Pesch, C., & Wielenga, C. (2020). Epistemic (in)justice and decolonisation in higher education: Experiences of a cross-site teaching project. *Acta Academica*, 52(1), 54-75. <https://doi.org/10.18820/24150479/aa52i1/SP4>
- King, R., & Raghuram, P. 2013. International student migration: Mapping the field and new research agendas. *Population, Space and Place*, 19(2), 127-137. <https://doi.org/10.1002/psp.1746>
- Mareque, M., De Prada, E., & Pino-Juste, M. (2020). Creativity and intercultural experiences: The impact of university international exchanges. Creativity. *Theories Research Applications*, 7(2), 321-345. <https://doi.org/10.2478/ctra-2020-0017>
- Martínez-Usarralde, M.-J., Pausá, J. M., & García-López, R. (2017). The ERASMUS experience and its capacitating potential: Analysis of adaptive capabilities. *International Journal of Educational Development*, 53, 101-109. <https://doi.org/10.1016/j.ijedudev.2017.01.001>
- Miller, A. L., Silberstein, S. M., & BrckaLorenz, A. (2021). Honors college faculty support for high-impact practice participation. *Journal for the Education of the Gifted*, 44(4), 1-26. <https://doi.org/10.1177/01623532211023599>
- Mohammed, A. S. (2021). Experience sharing on virtual COOP training to mechanical engineering students during the COVID-19 pandemic. *Sustainability*, 13(20), 11401. <https://doi.org/10.3390/su132011401>
- Monteiro, S. C., Caires, S. M. G., & Almeida, L. S. (2025). Teaching and learning among students' diversity: A project-based learning experience in higher education. *Revista Brasileira de Educação*, 30, e300051. <https://doi.org/10.1590/S1413-24782025300051>
- Patrício, M. T., Santos, P., Loureiro, P. M., & Horta, H. (2017). Faculty-exchange programs promoting change: Motivations, experiences, and influence of participants in the Carnegie Mellon University-Portugal Faculty Exchange Program. *Tertiary Education and Management*, 24(1), 1-18. <https://doi.org/10.1080/13583883.2017.1305440>
- Pearson, J., McNamara, K. E., & McMichael, C. (2022). International student mobility: Pacific Islander experiences of higher education in Australia. *Asia Pacific Viewpoint*, 63(2), 194-206. <https://doi.org/10.1111/apv.12342>
- Pherali, T., & Lewis, A. (2019). Developing global partnerships in higher education for peacebuilding: A strategy for pathways to impact. *Higher Education*, 78(4), 729-744. <https://doi.org/10.1007/s10734-019-00367-7>

Powell, P., & Walsh, A. (2018). Mutualising the university: Achieving community impact through an ecosystem. *International Review of Education*, 64(4), 563–583. <https://doi.org/10.1007/s11159-018-9739-5>

Rodríguez Marconi, D., Lapierre Acevedo, M., Serra, M., Zanetti Fontaine, L., Sanabria, C.-M., & Quiroz Almuna, H. (2023). Collaborative online international learning as a strategy for the development of transversal competences in higher education: An experience from the speech and language pathology degree. *Educación Médica*, 24, 100835. <https://doi.org/10.1016/j.edumed.2023.100835>

Ronnie, L. (2016). Academic capital and the world of work: Experiences of mature students. *Journal of Applied Research in Higher Education*, 8(2), 265–276. <https://doi.org/10.1108/JARHE-02-2015-0013>

Saidi, Z. A. & Abd Aziz, M. F. (2025). Empowering student leadership in higher education: A literature review on key competencies for effective leadership. *Journal of Institutional Research South East Asia*, 23(1), 376-399.

Saidi, Z.A., Asri, A.S., & Abd Aziz, M.F. (2025). Exploring the relationship between participation, needs, readiness, and planning in university student volunteering programs. *Journal of Institutional Research South East Asia*, 23(1), 401-417.

Sisavath, S. (2021). Benefits of studying abroad for graduate employability: Perspectives of exchange students from Lao universities. *Journal of International Students*, 11(3), 547-566. <https://doi.org/10.32674/jis.v11i3.2779>

Smith, H. A., & Segbers, T. (2018). The impact of transculturality on student experience of higher education. *Journal of Experiential Education*, 41(1), 75–89. <https://doi.org/10.1177/1053825917750406>

Søndergaard, M., Takita, F., & Van Rompay-Bartels, I. (2023). International students' perceptions towards their learning experience in an international network seminar in Japan: During and post the COVID-19 pandemic. *Sustainability*, 15(11), 8641.

Song, J. (2018). Creating world-class universities in China: Strategies and impacts at a renowned research university. *Higher Education*, 75(4), 729–742. <https://doi.org/10.1007/s10734-017-0167-4>

Valls-Figuera, R. G., Torrado-Fonseca, M., & Borràs, J. (2023). The impact of international student mobility on multicultural competence and career development: The case of students from Latin America and the Caribbean in Barcelona. *Education Sciences*, 13(9), 869. <https://doi.org/10.3390/educsci13090869>

Vande Berg, M., Paige, R. M., & Lou, K. H. (2009). Student learning abroad: What our students are learning, what they're not, and what we can do about it. Stylus Publishing.

Whetten, J., Fontenla, M., & Villa, K. (2018). Opportunities for higher education: The ten-year effects of conditional cash transfers on upper-secondary and tertiary enrollments. *Oxford Development Studies*, 47(2), 222-237.

Reverse Mentoring for Institutional Transformation in Higher Education: A Systematic Review of Adoption Patterns, Contextual Enablers, and Learning-Innovation Outcomes

ABSTRACT

Authors: Hongyi Huo, Faiq Aziz* & Mageswari Kunasegaran

Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

*Corresponding author:
mohdfaiq@upm.edu.my

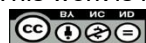
Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026:

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Reverse mentoring (RM) has emerged as a promising mechanism for intergenerational knowledge exchange in higher education, yet empirical evidence remains scattered across isolated studies without systematic synthesis. This review addresses the gap by examining how RM is adopted, implemented, and linked to organizational outcomes in higher education settings. Following PRISMA guidelines, searches of Scopus, Web of Science, and ERIC identified ten empirical studies—four from higher education and six from other sectors serving as comparative references. Thematic synthesis revealed four adoption patterns in higher education: governance-oriented, training-centred, curriculum-embedded, and community-partnered models. Six contextual enablers proved critical for successful implementation: governance sponsorship, open culture, dedicated resources, structured routines, external pressures, and evaluation mechanisms. Higher education institutions demonstrated particular strengths in evaluation culture and structural support through existing infrastructure such as teaching and learning centres. Three outcome themes emerged progressively: learning as foundational skill development, alignment as improved cross-level communication, and innovation as cumulative institutional change. Cross-sector comparison provided practical insights for universities, including quantitative outcome tracking from technology firms, process embedding from manufacturing, shadow committees from hospitality, and consortium resource-sharing from SMEs. These findings offer a coherent framework connecting RM's contextual conditions to its learning and innovation outcomes, offering practical guidance for university leaders.

Keywords: Reverse mentoring; Systematic literature review; Higher education; Organizational learning; Intergenerational collaboration; Institutional innovation

1. Introduction

Higher education institutions worldwide face unprecedented challenges arising from generational diversity and technological acceleration. As digital natives enter universities both as students and early-career staff, significant knowledge asymmetries emerge between those fluent in emerging technologies and senior administrators shaped by pre-digital experiences (Chaudhuri & Ghosh, 2012). Meanwhile, calls for participatory governance have intensified, with students increasingly demanding meaningful involvement in institutional decision-making rather than tokenistic representation (Matthews & Dollinger, 2023). These converging pressures provide conditions favouring reverse mentoring (RM), a developmental practice that inverts traditional hierarchical learning by positioning younger or less experienced individuals as mentors to senior colleagues. Unlike corporate settings where hierarchical barriers may be more rigid, universities possess unique characteristics—including cultures of inquiry, structured academic calendars, and existing mentoring infrastructure—that may either facilitate or complicate reverse mentoring implementation.

Reverse mentoring originated at General Electric in 1999, when Jack Welch paired junior employees with executives to transfer digital competencies upward (Marcinkus Murphy, 2012). The practice has since evolved beyond technological upskilling to encompass diversity awareness, cultural transformation, and inclusive leadership development (Marcinkus Murphy, 2012; Gordon, 2025). Within higher education, RM holds particular promise as a mechanism for bridging the student-leadership gap, enabling fresh perspectives to inform institutional strategy while developing student mentors' professional capabilities (Gündüz & Aksit, 2018; Raymond et al., 2021). Universities have deployed RM through teaching and learning centers, embedded it within academic courses, and extended it to community partnerships, showing RM's adaptability across contexts (Çevik et al., 2025; Juris et al., 2022).

Yet despite growing adoption, RM research remains fragmented. Empirical studies are scattered across sectors—corporate, healthcare, and higher education—without comparative synthesis (Dale, 2025). Theoretical frameworks are invoked selectively, with scholars drawing on social exchange theory to explain relational dynamics, social learning theory to account for knowledge transfer, and organizational learning theory to interpret institutional outcomes, but rarely integrating these perspectives (Gadomska-Lila, 2020). Consequently, understanding of when and why RM succeeds or fails remains incomplete. This gap carries practical consequences: university leaders either avoid adopting RM due to uncertainty about implementation requirements, or launch initiatives without adequate structural support, resulting in unsustainable programs that dissipate once initial enthusiasm fades.

To address these limitations, this study conducts a systematic literature review examining how RM is adopted, implemented, and linked to organizational outcomes in higher education settings. Following PRISMA guidelines, searches across Scopus, Web of Science, and ERIC identified ten empirical studies meeting inclusion criteria—four from higher education and six from other organizational sectors serving as comparative references. Through thematic synthesis, the review analyzes adoption patterns, contextual enablers, and learning-innovation outcomes, with higher education as the primary analytical focus and cross-sector comparison illuminating distinctive features and transferable insights.

Three research questions guide this inquiry:

- RQ1: What patterns of RM adoption characterize higher education institutions?
- RQ2: What contextual factors shape RM implementation in higher education settings?
- RQ3: What learning and innovation outcomes does RM produce in higher education?

Table 1 presents the mapping between research questions, analytical methods, and expected outputs.

Table 1. Research Question–Method–Output Mapping

Research Question	Data Source	Analysis Method	Output
RQ1: Adoption patterns	10 studies (4 HE primary, 6 other sectors comparison)	Cross-case classification	Table 4: Adopter typology with HE focus
RQ2: Contextual enablers	Coding of contextual factors across 10 studies	Cross-case thematic matrix	Figure 2: Enabler comparison by sector

RQ3: Outcomes	Evidence extraction with HE sub-analysis	Thematic synthesis	Figure 3: Outcome prevalence heatmap

This review contributes to the literature in three ways. First, it establishes a conceptual framework integrating contextual enablers and outcome dimensions specifically calibrated for higher education contexts. Second, it broadens the empirical foundation by systematically comparing RM implementation across sectors, revealing what universities can learn from corporate and community applications. Third, it provides evidence-based guidance for institutional leaders seeking to design sustainable RM initiatives that strengthen participatory governance and organizational learning.

2. Literature Review

2.1. Conceptual and Theoretical Foundations

Reverse mentoring refers to a developmental relationship in which a younger or less senior individual mentors a more experienced colleague, transferring knowledge and perspectives that flow counter to traditional organizational hierarchies (Chaudhuri & Ghosh, 2012). The concept emerged from General Electric’s initiative in 1999, when CEO Jack Welch paired 500 junior employees with senior executives to help leadership understand the emerging digital landscape (Marcinkus Murphy, 2012). What began as a technology transfer mechanism has since expanded to address diversity awareness, generational understanding, and cultural transformation across organizational settings (Marcinkus Murphy, 2012; Harvey & Allard, 2020).

Several theoretical perspectives inform the study of reverse mentoring. Social Exchange Theory explains how mentoring relationships are sustained through perceived mutual benefit rather than formal authority (Chaudhuri & Ghosh, 2012). In reverse mentoring, junior mentors offer technological expertise and contemporary perspectives, while senior mentees reciprocate with organizational insight, career guidance, and access to professional networks. The perception of fair exchange becomes critical to maintaining engagement, particularly when traditional power differentials are inverted (Rutti et al., 2013).

Social Learning Theory provides a complementary lens by emphasizing observational learning and behavioural modelling (Bandura, 1977). This perspective is further supported by evidence from higher education research emphasizing the role of specific leadership competencies in shaping effective developmental outcomes. Core competencies such as communication, interpersonal skills, critical thinking, emotional intelligence, and digital proficiency have been consistently identified as essential for leadership effectiveness in contemporary academic environments (Saidi & Abd Aziz, 2025). Reverse mentoring creates environments where both parties learn through demonstration, dialogue, and reflection—not merely through instruction. Senior leaders observe how younger colleagues navigate digital tools and contemporary workplace norms, while junior mentors gain exposure to strategic thinking and organizational politics. This bidirectional observation produces cognitive and behavioural changes that extend beyond the immediate mentoring relationship (Gong et al., 2024).

At the organizational level, reverse mentoring connects to Organizational Learning Theory, which conceptualizes institutions as entities capable of acquiring, interpreting, and embedding knowledge into routines and practices (Argyris & Schön, 1978). Reverse mentoring can function as a mechanism for “sensing” environmental shifts—particularly technological and generational changes—and translating individual learning into collective capability (Gupta et al., 2024). The Dynamic Capabilities Framework extends this perspective by positioning reverse mentoring as a means through which organizations sense emerging trends, seize opportunities for renewal, and transform their practices accordingly (Teece, 2007). When reverse mentoring is embedded in organizational routines rather than treated as a one-off initiative, it can contribute to sustained adaptive capacity.

Despite the availability of these theoretical tools, the literature tends to invoke them selectively. Studies emphasizing relational dynamics draw on social exchange; those focused on skill transfer lean toward social learning; and those examining institutional outcomes reference organizational learning or dynamic capabilities (Gadomska-Lila, 2020). Rare are attempts to integrate these perspectives into a unified framework that traces how individual exchanges aggregate into organizational outcomes. This theoretical fragmentation limits cumulative knowledge building and leaves practitioners without clear guidance on which mechanisms matter most under different conditions.

2.2. Evolution and Key Themes in RM Research

Over the past two decades, academic interest in reverse mentoring has grown and shifted in response to changing organizational priorities. During the early phase (2000–2010), research concentrated on technological and generational divides. Studies documented corporate experiments at firms like GE, IBM, and Procter & Gamble, noting improvements in executives' digital literacy but also challenges in sustaining engagement once initial novelty wore off (Chaudhuri & Ghosh, 2012). The dominant framing positioned reverse mentoring as a response to the “digital divide” between technology-fluent younger workers and senior leaders less comfortable with emerging tools.

The middle phase (2011–2020) expanded beyond technology transfer toward diversity and inclusion agendas. Reverse mentoring became incorporated into leadership development programs aimed at fostering empathy across lines of gender, ethnicity, and age (Marcinkus Murphy, 2012). Universities and public institutions began experimenting with reverse mentoring as a tool for inclusive governance, pairing students with administrators or pairing junior faculty with senior colleagues to surface marginalized perspectives (Zauchner-Studnicka, 2017). Methodological approaches also diversified during this period, with qualitative case studies and reflective interventions examining psychological safety, trust formation, and relational dynamics (Gadomska-Lila, 2020).

Recent research (2021–present) increasingly frames reverse mentoring as an organizational learning and innovation mechanism rather than merely a means of role reversal. Researchers such as Gupta et al. (2024) and Li et al. (2025) examine how reverse mentoring contributes to strategic renewal through cross-level knowledge integration. The analytical focus has shifted from interpersonal learning to system-level processes, including capability building and cultural agility. Garg et al. (2021) empirically linked reverse mentoring practices with innovative behavior in technology firms, demonstrating that learning engagement mediates the relationship between mentoring participation and innovation outcomes. This recent turn positions reverse mentoring not as a peripheral HR intervention but as a potential driver of organizational transformation.

2.3. Reverse Mentoring in Higher Education

Higher education institutions have characteristics that suit them to reverse mentoring while also presenting distinct challenges. As knowledge-intensive organizations, universities are fundamentally oriented toward learning, inquiry, and the exchange of ideas—values that align naturally with mentoring practices (Chun & Evans, 2021; Garvey et al., 2017). Students occupy a unique position as “insider-outsiders”: they are embedded within the institution yet bring external perspectives shaped by generational experiences, technological fluency, and diverse backgrounds that senior administrators may lack direct access to. These features align with what organizational theorists term “professional bureaucracies”—structures coordinated through expertise standardization rather than hierarchical supervision, where change proceeds through gradual professional evolution rather than top-down mandate (Mintzberg, 2024).

Several features of higher education facilitate reverse mentoring implementation. Teaching and learning centers provide natural institutional homes for mentoring initiatives, offering training infrastructure, coordination capacity, and evaluation expertise (Çevik et al., 2025). The academic calendar creates structured rhythms—semesters, terms, and academic years—that can anchor mentoring cycles and provide natural endpoints for assessment. Perhaps most importantly, universities possess deeply rooted cultures of reflection, feedback, and assessment that support the evaluative practices essential for program improvement and sustainability.

Yet higher education also presents obstacles. Academic hierarchies, while perhaps less rigid than corporate structures, still shape interaction patterns between students and senior leaders. Administrators may feel uncomfortable receiving guidance from students they are accustomed to directing. Students, for their part, may lack confidence to offer candid feedback to figures of institutional authority (Yurtseven Avci & Eren, 2023). Successful reverse mentoring in universities therefore requires deliberate cultivation of psychological safety and explicit institutional endorsement to legitimate the inverted relationship (Gündüz & Aksit, 2018).

The applications of reverse mentoring in higher education span governance, teaching, and community engagement. Governance-oriented programs pair students directly with senior administrators to inform strategic decisions and surface student perspectives on institutional priorities. Curriculum-embedded models integrate reverse mentoring into academic courses, providing students with structured learning experiences while connecting organizations with fresh viewpoints (Raymond et al., 2021). Community-partnered initiatives extend reverse mentoring beyond campus boundaries, linking university students with older adults or community members in intergenerational learning exchanges (Juris et al., 2022). Each model reflects different institutional priorities and draws on different governance structures for support.

2.4. Research Gaps

Despite growing interest, several gaps constrain current understanding of reverse mentoring in higher education. First, theoretical integration remains limited. Scholars invoke multiple frameworks—social exchange, social learning, organizational learning, dynamic capabilities—but rarely articulate how these perspectives connect or specify conditions under which one mechanism predominates over another. For instance, no study has examined how social exchange dynamics at the dyadic level translate into organizational learning at the institutional level. Second, empirical evidence is fragmented across isolated case studies, making it difficult to identify patterns that hold across contexts or to distinguish findings that reflect unique circumstances from those with broader applicability. Existing studies have also tended to examine leadership competencies and experiential engagement in isolation, with limited integration of how these elements interact within structured developmental practices such as reverse mentoring (Saidi & Abd Aziz; 2025; Saidi et al., 20205). Third, the relationship between contextual conditions and outcomes remains underspecified. While governance support, cultural openness, and resource availability are frequently cited as important, the field lacks a systematic framework connecting specific enablers to particular outcomes. This review responds to Chaudhuri et al.'s (2022) call for research examining reverse mentoring beyond corporate contexts. By centering higher education and developing a sector-calibrated enabler-outcome framework, it addresses their observation that existing studies “rarely articulate how theoretical perspectives connect” (p. 472).

2.5. Conceptual Framework

Drawing on the theoretical perspectives reviewed above, this study proposes an integrative framework to guide analysis as shown in Figure 1. The framework identifies three categories of contextual enablers that shape reverse mentoring implementation.

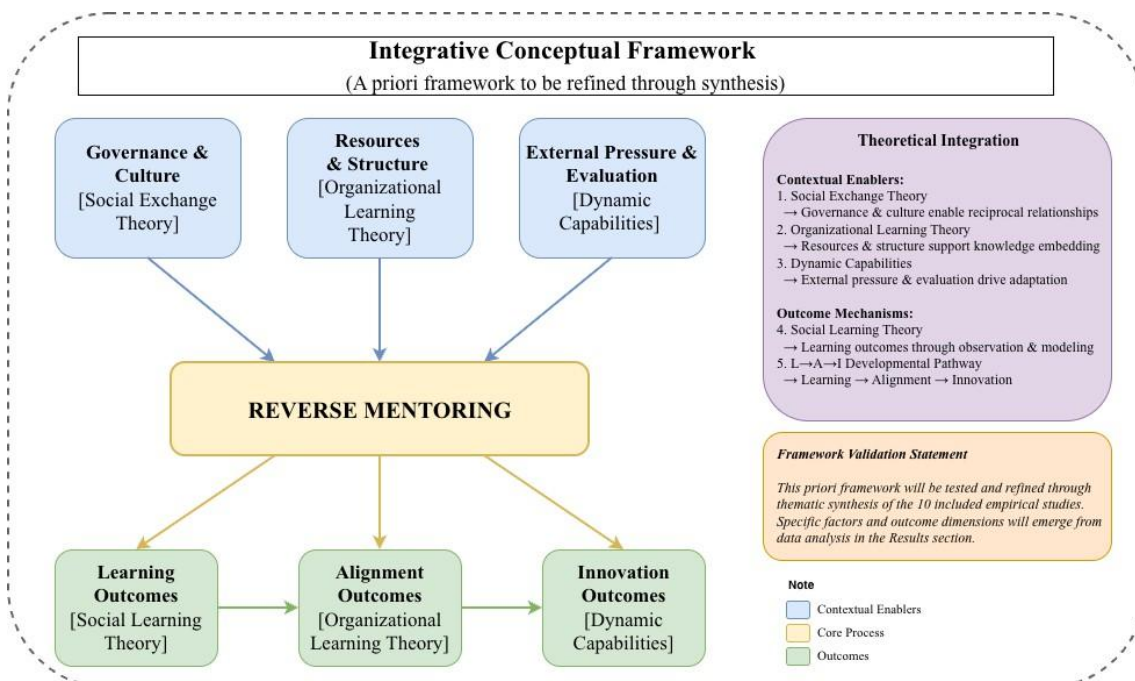


Figure 1. Integrative Conceptual Framework (Prior Analytical Framework)

The first category, governance and culture, draws on Social Exchange Theory (Chaudhuri & Ghosh, 2012). When participants perceive fair reciprocity, mentoring relationships are more likely to form and sustain. The second category, resources and structure, applies Organizational Learning Theory (Argyris & Schön, 1978). Adequate support systems help convert individual insights into collective knowledge. The third category, external pressure and evaluation, builds on Dynamic Capabilities (Teece, 2007). Environmental demands and feedback mechanisms prompt organizations to adapt.

The framework also proposes that reverse mentoring outcomes may follow a developmental sequence. Learning outcomes—grounded in Social Learning Theory (Bandura, 1977)—are theorized to form the foundation through observation and modelling. Alignment outcomes may build upon learning as knowledge spreads across levels, reflecting Organizational Learning processes. Innovation outcomes are hypothesized to emerge when learning and alignment become embedded in organizational routines, enabled by Dynamic Capabilities. This proposed L→A→I sequence draws on Crossan et al.'s (1999) 4I framework, where individual insights aggregate into institutional change. However, this sequence represents a theoretical proposition requiring longitudinal validation; cross-sectional studies cannot establish temporal ordering.

This framework serves as an analytical scaffold for the synthesis that follows. The three enabler categories will be refined into specific factors through analysis of the included studies. The outcome dimensions will also be elaborated based on empirical findings. This allows data to validate, challenge, or extend the proposed relationships.

3. Methodology

This study adopted a systematic literature review approach following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The review aimed to synthesize empirical evidence on reverse mentoring adoption, implementation, and outcomes, with particular attention to higher education contexts.

3.1. Identification

Literature searches were conducted across three databases: Scopus, Web of Science, and ERIC on September 15th, 2025. They were selected for their comprehensive coverage of management, organizational studies, education, and human resource development. The search period spanned January 2007 to October 2025, as structured reverse mentoring programs in organizational and educational settings began appearing in the academic literature around 2007. Table 2 presents the search strings employed across databases and initial retrieval results.

Table 2. Search Strings

Database	Field Used	Search String
Web of Science	TS	TS = (("reverse mentoring*" OR "upward mentoring*") AND ("enterprise*" OR "company" OR "organization*"))
Scopus	TITLE-ABS-KEY	TITLE-ABS-KEY(("reverse mentoring*" OR "upward mentoring*") AND ("enterprise*" OR "company" OR "organization*"))
ERIC	ANYWHERE	TX("reverse mentoring*" OR "upward mentoring*") AND ("organization*" OR "enterprise*" OR "workplace")

3.2. Screening

All retrieved articles (n=35) were screened using explicit inclusion and exclusion criteria to ensure the relevance, accessibility, and methodological quality of the selected studies. The criteria applied during the screening process are summarized in Table 3.

Table 3. Inclusion and Exclusion Criteria

Criteria Type	Inclusion Criteria	Exclusion Criteria
Publication Type	Peer-reviewed journal articles: empirical studies	Non-peer-reviewed sources, book chapters, review articles, or grey literature

Language	English	Non-English
Time Frame	January 2007 – October 2025	Articles published before 2007
Research Focus	Studies examining reverse mentoring (RM) in organizational contexts	Studies focused only on traditional mentoring, unrelated leadership training, or other non-RM programs.
Accessibility	Full-text, open-access	Non-open access, abstract-only records, or inaccessible documents

3.2.1. Cross-Sector Inclusion Rationale

The decision to include non-higher education studies alongside university research reflects the concept's corporate origins (Marcinkus Murphy, 2012) and the scarcity of higher education empirical research. Cross-sector comparison enables identification of context-dependent versus transferable enablers—a strategy aligned with realist synthesis principles that emphasize understanding “what works, for whom, under what circumstances” (Pawson, 2006, p. 25).

3.2.2. Open-Access Inclusion Rationale and Bias Assessment

The decision to include only open-access peer-reviewed articles was made for two reasons: first, to ensure full-text accessibility for thorough analysis and potential replication; second, to align with principles of research transparency increasingly emphasized in educational scholarship. However, this choice introduces potential limitations that require explicit acknowledgment. A supplementary search without the open-access filter identified approximately 12 additional empirical studies that could not be accessed for full-text review. Table 4 documents notable excluded studies and their characteristics.

Table 4. Notable Paywalled Studies Excluded Due to Open-Access Filter

Study	Journal	Sector	Country	Focus
Marcinkus Murphy (2012)	<i>Human Resource Management</i>	Corporate	USA	Foundational RM framework
Garg, Murphy & Singh (2021)	<i>Career Development International</i>	Technology	India	RM, work engagement, innovation
Gadomska-Lila (2020)	<i>Journal of Organizational Change Management</i>	Multiple	Poland	RM intergenerational relationships
Kaše, Saksida & Mihelič (2019)	<i>Human Resource Management</i>	Corporate	Slovenia	Skill development in RM
Harvey et al. (2009)	<i>Int'l J. of Human Resource Management</i>	Global corporations	USA	Reverse and reciprocal mentoring
Upadhyay & Singh (2024)	<i>Int'l J. of Organizational Analysis</i>	Technology	India	Leader humility and RM
O'Connor et al. (2025)	<i>British Educational Research Journal</i>	Higher Education	UK	Student-leadership RM for EDI

Note: Studies identified through supplementary searches of Scopus and Web of Science without open-access restriction.

The paywalled studies are concentrated in European subscription journals and include quantitative designs underrepresented in our sample. This suggests the open-access filter may introduce geographic bias (under-representing European research), journal-type bias (excluding premier management journals such as *Academy of Management Journal* and *Organization Science*), and methodological bias (under-representing quantitative studies). The cross-sector comparison in this review therefore contrasts open-access-available research rather than necessarily representative practice from each sector.

Despite these limitations, the open-access filter was retained because it enables full reproducibility by researchers without institutional access barriers. Based on abstract review, the excluded paywalled studies do not appear to contradict the core thematic patterns identified regarding adoption models, enabler categories, and outcome dimensions. Future reviews with broader institutional access should expand searches to include

subscription journals, particularly for strengthening cross-sector comparisons and quantitative evidence synthesis.

3.3. Included

3.3.1. Retrieval and Deduplication

Literature searches were conducted on September 15, 2025. Table 5 presents the initial retrieval counts by database and the deduplication process.

Table 5. Database Retrieval and Deduplication

Stage	Scopus	Web of Science	ERIC	Total
Initial records retrieved	11	11	13	35
Duplicates removed (Mendeley)	2	1	1	4
Duplicates removed (manual check)	0	0	1	1
Records after deduplication	9	10	11	30

Note: Mendeley’s duplicate detection function identified 4 duplicate entries. Manual review identified 1 additional duplicate with variant author name spellings.

3.3.2. Screening Workflow

Title, abstract, and keyword screening was conducted by the first author on the 30 deduplicated records. Fifteen articles were excluded at this stage for the following reasons: not an empirical study ($n = 5$), not focused on reverse mentoring ($n = 6$), and not in an organizational context ($n = 4$). The second author independently screened a random sample of 12 records (40%) to assess reliability; agreement was reached on 11 of 12 records (92%), with the single discrepancy resolved through discussion.

The remaining 15 articles proceeded to full-text eligibility assessment. Full-text versions were retrieved and evaluated against inclusion and exclusion criteria by the first author. The second author independently reviewed all 15 full-texts. Five articles were excluded at this stage: no open-access full text available ($n = 2$), not a peer-reviewed article ($n = 1$), and insufficient empirical data ($n = 2$). Disagreements on 2 records were resolved through discussion until consensus was reached. Table 6 presents the complete exclusion reasons across both screening stages.

Table 6. Full-Text Exclusion Reasons

Stage	Exclusion Reason	Count
Title/Abstract Screening	Not empirical study	5
	Not focused on RM	6
	Not organizational context	4
	<i>Subtotal</i>	<i>15</i>
Full-Text Screening	No open-access full text	2
	Not peer-reviewed article	1
	Insufficient empirical data	2
	<i>Subtotal</i>	<i>5</i>
Total Excluded		20
Studies Included		10

After this assessment, ten studies met all requirements and were retained for synthesis. Figure 2 illustrates the selection process following PRISMA conventions.

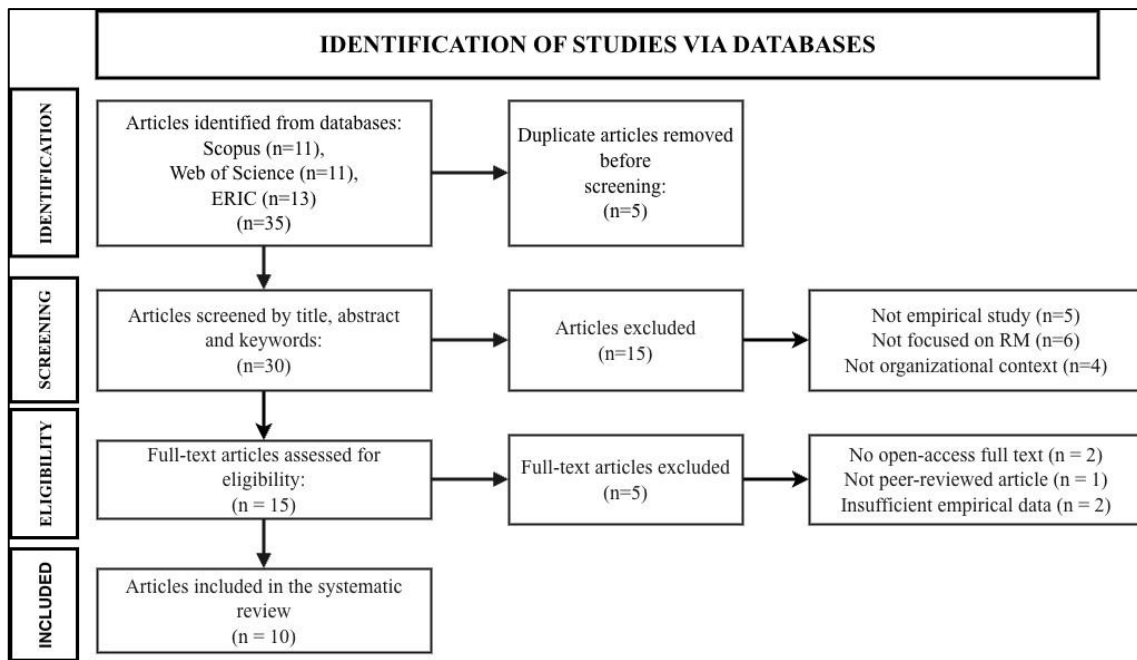


Figure 2. PRISMA Flow Diagram Adapted from Page et al. (2021)

3.4. Quality Appraisal

The methodological quality of included studies was assessed using the Crowe Critical Appraisal Tool (CCAT), which evaluates eight dimensions: preliminaries, introduction, design, sampling, data collection, ethical matters, results, and discussion (Crowe & Sheppard, 2011). Each dimension receives a score from 0 to 5, yielding a maximum total of 40 points, as shown in Appendix A. All ten studies achieved scores indicating acceptable methodological quality. Following Pawson (2006) and Dixon-Woods et al. (2006), no studies were excluded solely on quality grounds. This inclusive approach is appropriate for reviews seeking to understand mechanisms and contextual patterns rather than to calculate precise effect sizes, as excluding lower-scoring studies may reduce conceptual diversity.

3.5. Data Extraction and Coding

As can be seen in Table 7, a structured extraction protocol captured key information from each study: bibliographic details, geographic context, organizational sector, research design, sample characteristics, reverse mentoring program features, contextual factors, and reported outcomes.

Table 7: Data Extraction Fields

Category	Extracted Information
Study identification	Authors, year, country/region
Context	Organizational sector, institutional setting
Methods	Research design, sample size, data sources
RM program	Program type, duration, pairing approach
Contextual factors	Governance, culture, resources, structure, external pressures, evaluation
Outcomes	Learning, alignment, innovation—with supporting evidence

The coding process followed a hybrid approach combining deductive and inductive elements. The initial coding frame drew on categories identified in the literature review—contextual enablers and outcome types—while remaining open to themes emerging from the data. Coding was conducted by the first author using NVivo 14. The co-author independently reviewed coding of all 4 Higher Education studies (100%) and 2 randomly selected non-HE studies (33%); discrepancies were resolved through discussion until consensus was reached.

Thematic synthesis followed the approach outlined by Thomas and Harden (2008). This involved line-by-line coding of relevant passages, development of descriptive themes capturing patterns within the data, and

generation of analytical themes that extend beyond the content of individual studies to address the research questions. Coding dictionary is shown in Appendix B.

To assess coding reliability, the second author independently coded all 4 higher education studies (100% of HE corpus) and 2 randomly selected non-HE studies (33% of non-HE corpus), totalling 6 studies (60% of full corpus). Initial agreement on contextual enabler coding (A1–C2) reached 85%, and agreement on outcome coding (O-L, O-A, O-I) reached 90%. Discrepancies primarily involved distinguishing “Partial” from “Y” coding for Resources (B1), where some studies mentioned resources without specifying dedicated allocation, and whether certain impacts constituted Alignment versus Learning when studies described both communication improvement and skill development. All discrepancies were resolved through discussion until consensus was reached, with coding rules refined iteratively to improve clarity.

3.6. Summary of Included Studies

Table 8 provides an overview of the ten studies included in the synthesis, organized by sector. The complete data extraction table is available as supplementary materials.

Table 8. Summary of the Selected Studies

Study	Aim	Findings
Lužar et al. (2023)	To investigate impact of intergenerational differences on knowledge transfer among Slovenian engineers	Intergenerational differences (esp. tech skills) exist; RM & supportive culture improve knowledge transfer
Gupta et al. (2024)	To develop digital innovation and transformation capabilities in a large company (case study)	Key mechanisms (digital champions, council, core team, reverse mentoring) foster cultural change; coercive & mimetic pressures drive workforce upskilling
Chen (2013)	To explore reverse mentoring’s effect on traditional mentoring functions (career development, psychosocial support, role modeling)	Gen X/Y juniors have advanced IT skills, motivation, innovative thinking; reverse mentoring retains traditional functions (career, psychosocial, role-modeling)
Gündüz & Aksit (2018)	To explore student–university president reverse mentoring at Maltepe University and its significance for higher education	Highlights the importance of proper training for reverse mentors. Providing a Student–president reverse mentoring benefits universities broadly and helps dispel negative stereotyping in academia
Çevik et al. (2025)	To investigate effectiveness of a video-based online training program for student mentors in technology mentoring	Training significantly improved student mentors’ skills; mentors felt better prepared; mentees rated mentors highly in communication, enthusiasm, organization, planning
Schwartz et al. (2022)	To develop and pilot-test a new reverse mentoring model for SMEs to overcome innovation constraints	SMEs generated innovative ideas and adopted open innovation via this model; pilot showed feasibility and efficacy
Raymond et al. (2021)	To implement a student–employee reverse mentoring program in a marketing course (with a Fortune 500 company) to boost students’ career readiness	Reverse mentoring helped executives with knowledge continuity plans and improved students’ competencies (critical thinking, teamwork, professionalism, communication)
Li et al. (2025)	To examine how reverse mentoring influences employees’ innovative behavior and the roles of challenge stress and continuous learning	Reverse mentoring significantly boosts employees’ innovative behavior; effect is mediated by young mentors’ challenge-stress appraisal and continuous learning
Juris et al. (2022)	To use a virtual intergenerational reverse mentoring program (Cyber-Seniors) to reduce social isolation and loneliness in older adults	Older adult participants reported decreased loneliness post-program (learned new ways to connect); student mentors’ loneliness was unchanged
Cismaru & Iunius (2020)	To confirm reverse mentoring as a valuable talent management approach to bridge generational gaps in the hospitality industry	Reverse mentoring shows strong innovative potential in hospitality; employees have high interest but doubt management will implement it

4. Results and Findings

RQ1: What Patterns of RM Adoption Characterize Higher Education Institutions?

Ten studies were analyzed to identify organizational types that have adopted reverse mentoring. Classification was based on which entity designs, runs, and evaluates the program rather than on participant numbers. Five adopter types emerged: higher education (4 studies), technology enterprises (2), large manufacturing (2), hospitality (1), and SMEs (1). Table 9 summarizes these patterns.

Table 9. Types of Organizations Adopting RM

Type of Organization	No.	Article	Country	Governance Locus
Higher Education	4	Gündüz & Aksit (2018); Çevik et al. (2025); Raymond et al. (2021); Juris et al. (2022)	Turkey (2), USA (2)	Academic units, TLCs, governance offices
Technology Enterprises	2	Chen (2013); Li et al. (2025)	Taiwan, China	HR/L&D departments
Large Manufacturing	2	Gupta et al. (2024); Lužar et al. (2023)	India, Slovenia	Digital transformation offices
Hospitality & Services	1	Cismaru & Iunius (2020)	Switzerland, Romania	HR talent pipelines
SMEs	1	Schwartz et al. (2022)	Brazil	External industry federation

Geographic distribution merits attention. The corpus spans nine countries: Turkey (2), USA (2), Taiwan (1), China (1), India (1), Slovenia (1), Romania (1), and Brazil (1). Western and East Asian contexts predominate; no studies examined African, Middle Eastern, or Southeast Asian higher education systems, limiting generalizability to institutions operating under different governance traditions.

The four higher education studies showed distinct adoption patterns. Gündüz and Aksit (2018) documented a governance-oriented model at Maltepe University where students mentored senior executives including the president, with the program anchored in top leadership. Çevik et al. (2025) reported a training-centred model at Hacettepe University's Teaching and Learning Center, emphasizing structured mentor preparation through online courses. Raymond et al. (2021) described a curriculum-embedded model where U.S. marketing students mentored Fortune 500 employees for course credit. Juris et al. (2022) examined a community-partnered model linking Appalachian State University students with older adults through a regional aging agency. These four patterns are not mutually exclusive; institutions may combine elements of multiple models, as when a training-centred program also embeds curriculum components.

Despite these differences, the higher education cases shared common features: all operated through established university units rather than as standalone projects; all aligned with academic calendars; and all incorporated formal evaluation through surveys, journals, or course assessment. Students served as “insider-outsiders”—embedded in the institution yet bringing fresh generational perspectives that senior staff might otherwise miss.

Cross-sector comparison revealed important differences. Technology firms (Chen, 2013; Li et al., 2025) embedded reverse mentoring in HR systems with project-based timelines and performance-linked evaluation. Manufacturing companies (Gupta et al., 2024; Lužar et al., 2023) connected the practice to digital transformation agendas. Hospitality (Cismaru & Iunius, 2020) used “shadow committee” arrangements where junior staff observed management decisions. SMEs (Schwartz et al., 2022) relied on external federation support to compensate for limited internal resources.

Table 10. Cross-Sector Insights for Higher Education

Source Sector	Distinctive Feature	Implication for HE

Technology Enterprises	Integration with HR systems, rapid iteration	Embed RM in faculty development; establish quick feedback loops
Large Manufacturing	Alignment with strategic transformation	Link RM to university priorities like digital learning initiatives
Hospitality & Services	Shadow board mechanisms	Create student shadow committees beyond traditional representation
SMEs	Federation-based resource sharing	Form regional consortia to share training and evaluation tools

These comparisons suggest that higher education institutions are well-positioned for reverse mentoring given their existing infrastructure, evaluation culture, and semester-based rhythms. Table 10 presents how universities can learn from other sectors: technology firms' agile iteration, manufacturing's strategic alignment, hospitality's shadow governance mechanisms, and SMEs' consortium approaches all offer transferable lessons for strengthening university-based programs.

RQ2: What Contextual Factors Shape RM Implementation in Higher Education Settings?

To identify conditions that shape reverse mentoring implementation, six contextual factors were coded across the ten studies and grouped into three pairs: governance and culture (A1–A2), resources and structure (B1–B2), and external pressures and evaluation (C1–C2). Figure 3 displays the prevalence of each factor by organization type.

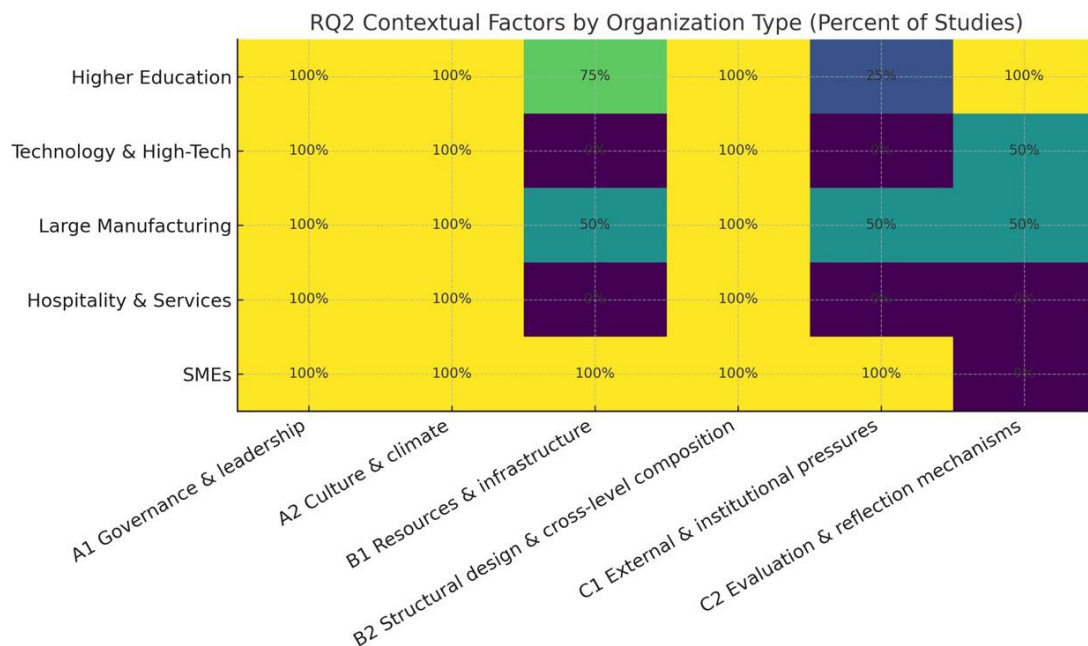


Figure 3. Contextual Factors by Organization Type

4.1. Governance and Culture (A1–A2)

Governance and culture together determine whether reverse mentoring moves from informal exchange to established practice. All ten studies reported governance mechanisms (A1) and cultural conditions (A2), confirming these as foundational requirements across sectors.

In higher education, governance takes distinctive forms. At Maltepe University, the president's direct involvement legitimized student mentors and signaled institutional commitment (Gündüz & Aksit, 2018). Hacettepe University's Teaching and Learning Center provided a formal home for mentor training and program coordination (Çevik et al., 2025). The U.S. marketing course operated through academic governance structures that ensured faculty oversight and curricular alignment (Raymond et al., 2021), while the Appalachian State initiative was anchored in community engagement offices (Juris et al., 2022). These arrangements share one

feature: reverse mentoring succeeds when it has a clear institutional owner rather than floating as an orphan initiative.

Cultural openness complements formal governance. Universities reported deliberate efforts to cultivate psychological safety—conditions where students feel comfortable offering candid feedback to senior staff and where administrators demonstrate genuine receptivity to challenge. Manufacturing firms similarly emphasized trust-building before expecting upward feedback (Lužar et al., 2023), suggesting that cultural preparation is necessary regardless of sector. The hospitality study noted that service-oriented norms supporting customer responsiveness extended to internal relationships, creating natural receptivity to junior input (Cismaru & Iunius, 2020).

4.2. Resources and Structure (B1–B2)

Resources and structural design provide the operational foundation for reverse mentoring. Here, notable differences emerged across sectors. Higher education showed strong structural design (B2 = 100%) but more varied resource investment (B1 = 75%). Three of four university studies reported dedicated training, platforms, or funding, while one relied primarily on existing course infrastructure without additional resource allocation.

Universities benefit from built-in advantages: teaching and learning centers offer training expertise, learning management systems provide communication platforms, and academic calendars create natural program rhythms. Çevik et al. (2025) developed a structured online course for mentor preparation, while Raymond et al. (2021) leveraged course assignments and reflection requirements. These existing assets reduce the marginal cost of reverse mentoring compared to organizations that must build infrastructure from scratch.

Technology and hospitality sectors showed a different pattern—strong structural design but minimal dedicated resources (B1 = 0%). These organizations embedded reverse mentoring within existing HR or team processes rather than creating separate resource streams. Manufacturing took a middle path, with half of studies reporting dedicated resources linked to digital transformation budgets (Gupta et al., 2024). SMEs compensated for internal resource limitations through external federation support, with FIEC providing training and coordination that individual small firms could not sustain (Schwartz et al., 2022).

4.3. External Pressures and Evaluation (C1–C2)

External pressures (C1) and evaluation mechanisms (C2) showed the greatest variation across sectors. Higher education reported the strongest evaluation practices (C2 = 100%) but the weakest external pressure drivers (C1 = 25%). Such a pattern reflects the academic culture of systematic assessment: all four university studies employed surveys, reflective journals, pre-post measures, or course-based evaluation. Universities evaluate because assessment is embedded in academic practice, not because external forces demand it.

External pressures operated differently across sectors. SMEs showed complete external pressure coverage (C1 = 100%), as the Brazilian pilot was explicitly driven by regional development policy and industry federation initiatives (Schwartz et al., 2022). Manufacturing firms responded to digital transformation imperatives and workforce demographic shifts (Gupta et al., 2024; Lužar et al., 2023). In contrast, university programs emerged more from internal pedagogical or governance goals than from external mandates—only one higher education study mentioned external drivers such as community partnership expectations (Juris et al., 2022).

4.4. Cross-Sector Insights for Higher Education

The cross-sector comparison suggests several lessons for universities. First, higher education's evaluation strength is a genuine asset—other sectors might learn from university assessment practices. Second, the relative weakness on external pressures may reflect autonomy: universities can implement reverse mentoring based on educational rationale rather than waiting for external mandates. Third, the resource variation across university studies indicates that reverse mentoring need not require substantial new investment if existing infrastructure is leveraged creatively.

From manufacturing and SMEs, universities can learn to connect reverse mentoring more explicitly to strategic priorities. When programs align with institutional goals—digital pedagogy, student voice in governance,

inclusive leadership—they gain legitimacy and resource support beyond what standalone initiatives typically attract. The SME consortium model also suggests that smaller institutions lacking teaching and learning centers might collaborate regionally to share training resources and evaluation tools.

RQ3: What Learning and Innovation Outcomes Does RM Produce in Higher Education?

Three outcome themes were coded across the ten studies: learning (O-L), alignment (O-A), and innovation (O-I). Learning refers to skill development, knowledge acquisition, and perspective change at individual or group levels. Alignment captures improved communication and coordination across hierarchical or generational boundaries. Innovation encompasses process improvements, institutional changes, and creative outputs at the organizational level. Figure 4 displays outcome prevalence by sector.

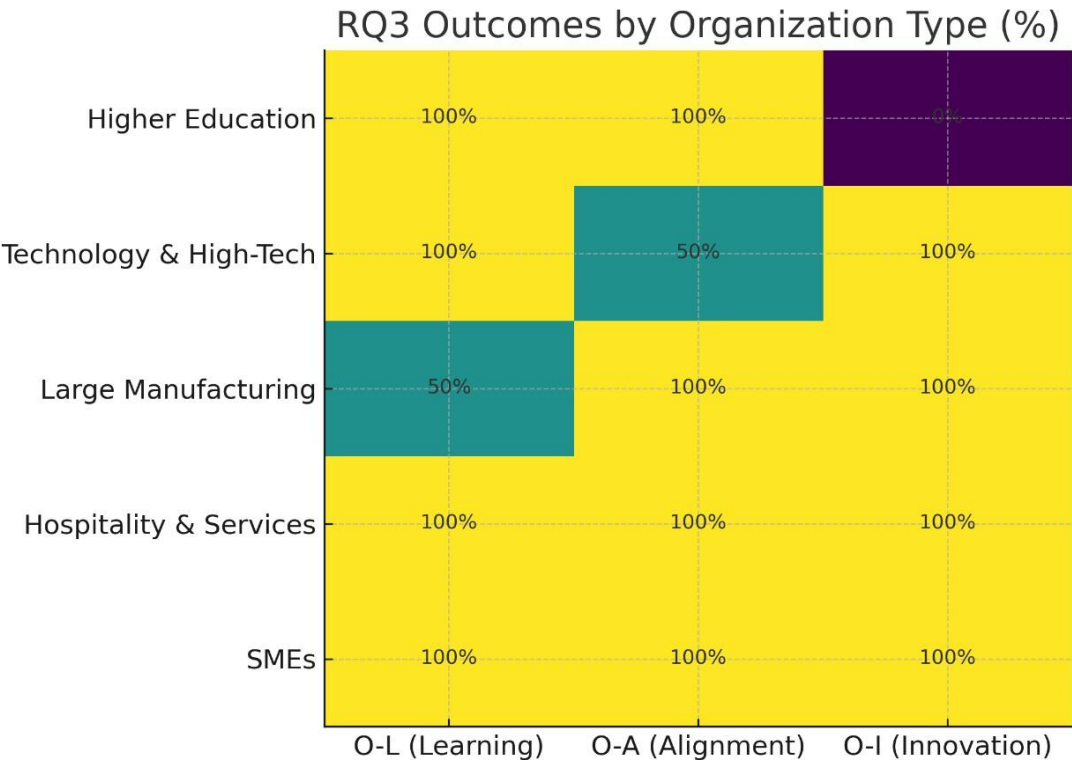


Figure 4. Outcomes Heatmap (O-L/O-A/O-I by type, %)

4.5. Learning Outcomes (O-L)

Learning was the most consistently reported outcome across all sectors. All four higher education studies documented learning gains, making O-L the strongest finding in university contexts. Student mentors reported growth in communication skills, organizational awareness, and professional confidence. At Hacettepe University, the structured training course strengthened students’ mentoring competence and reflective abilities (Çevik et al., 2025). Raymond et al. (2021) found that marketing students developed career readiness through their mentoring experience with Fortune 500 employees. The intergenerational program at Appalachian State enhanced students’ understanding of aging and community engagement (Juris et al., 2022).

Senior mentees also learned. At Maltepe University, administrators gained direct exposure to student perspectives they would not encounter through formal channels (Gündüz & Aksit, 2018). Older adults in the community partnership developed digital literacy skills through patient guidance from student mentors (Juris et al., 2022). All four studies emphasized bidirectional learning—both parties gained rather than knowledge flowing in only one direction.

Other sectors showed similar patterns. Technology firms reported mutual learning around digital tools and organizational context (Chen, 2013; Li et al., 2025). Manufacturing studies documented knowledge exchange

between generations of engineers (Gupta et al., 2024; Lužar et al., 2023). Learning appears to be the foundational outcome that reverse mentoring reliably produces regardless of setting.

4.6. Alignment Outcomes (O-A)

Alignment—improved cross-level communication and understanding—was reported in all four higher education studies, matching the learning outcome in prevalence. The Maltepe University case demonstrated how student-president mentoring reduced informational distance between students and top leadership (Gündüz & Aksit, 2018). Rather than relying on filtered reports through administrative layers, the president received direct student input on institutional priorities. This created feedback loops that became embedded in governance routines. Similarly, the intergenerational program built social connections that reduced isolation among older adults while helping students understand community needs (Juris et al., 2022).

In technology and manufacturing, alignment took different forms. Chen (2013) reported improved coordination between junior and senior engineers working on shared projects. Manufacturing studies showed how reverse mentoring facilitated knowledge integration across organizational levels, with younger staff helping translate new technologies for experienced colleagues and vice versa (Gupta et al., 2024; Lužar et al., 2023). The hospitality study connected alignment to vertical communication channels and shadow committee participation (Cismaru & Iunius, 2020).

4.7. Innovation Outcomes (O-I)

Innovation outcomes showed the starkest sectoral difference. None of the four higher education studies reported explicit innovation outcomes (O-I = 0%), while technology, manufacturing, hospitality, and SMEs all documented innovation at 100%. This finding requires interpretation rather than concern. Higher education studies focused on learning and governance goals—developing student capabilities, informing leadership decisions, building community connections—rather than commercial innovation or process improvement. The absence of reported innovation does not mean reverse mentoring fails in universities; it reflects different institutional priorities and outcome definitions.

In contrast, corporate and SME contexts explicitly linked reverse mentoring to innovation. Li et al. (2025) used a two-wave survey to demonstrate that reverse mentoring predicted innovative behavior among technology employees, with learning engagement serving as a mediating mechanism. Manufacturing studies connected the practice to digital transformation and process optimization (Gupta et al., 2024). The SME pilot documented small-scale innovations in workflow and customer engagement (Schwartz et al., 2022).

4.8. Cross-Sector Insights for Higher Education

The outcome comparison suggests that higher education excels at what it prioritizes—learning and relational alignment—but may underexplore reverse mentoring's innovation potential. Several insights emerge from other sectors. From technology, the Li et al. (2025) study offers a methodological lesson: innovation outcomes can be measured through surveys tracking creative behavior and idea generation. Universities could adapt such instruments to assess whether reverse mentoring stimulates teaching innovation, curricular creativity, or novel governance approaches. The finding that learning engagement mediates innovation suggests universities should design programs that maximize active participation rather than passive information exchange.

From manufacturing, the connection between reverse mentoring and strategic transformation (Gupta et al., 2024) suggests universities might frame the practice as supporting institutional change initiatives. When reverse mentoring aligns with goals like digital pedagogy or student experience improvement, innovation outcomes become more visible and measurable. From hospitality, the shadow committee model (Cismaru & Iunius, 2020) shows how junior voice can inform operational decisions. Universities might examine whether student mentors' insights lead to tangible changes in policies, services, or practices—a form of innovation that current studies may capture under alignment rather than innovation.

The pattern across outcomes appears consistent with the developmental sequence proposed in the conceptual framework: learning forms the foundation, alignment builds upon learning through improved relationships, and innovation may emerge when learning and alignment become embedded in organizational routines. However, this interpretation requires qualification. Only Li et al. (2025) employed a longitudinal design capable of establishing temporal sequencing; the remaining nine studies used cross-sectional approaches that document outcome presence but cannot confirm causal ordering. Alternative sequences remain possible—for instance, external pressure might trigger innovation attempts that subsequently drive learning, or strong alignment might enable rapid progression from learning to innovation. The L→A→I sequence should therefore be understood as a theoretical proposition consistent with the evidence rather than an empirically validated pathway. Higher education demonstrates learning and alignment clearly; the innovation stage may require longer timeframes, different measurement approaches, or more deliberate attention to document.

Table 11 summarizes the cross-study evidence patterns for contextual enablers and outcomes, with exemplar evidence from included studies demonstrating how codes were derived from primary texts.

Table 11. Evidence Patterns with Exemplar Quotations

Theme	Prevalence	Exemplar Evidence	Source
Contextual Enablers			
A1: Governance sponsorship	10/10 (100%)	"The president's direct involvement in the mentoring program signaled institutional commitment and legitimized student mentors in their advisory roles."	Gündüz & Akşit (2018)
A2: Open culture	10/10 (100%)	"Psychological safety was deliberately cultivated through orientation sessions where administrators expressed genuine openness to student feedback."	Çevik et al. (2025)
B1: Dedicated resources	7/10 (70%)	"The Teaching and Learning Center allocated staff time, developed training materials, and provided a learning management system platform for mentor-mentee communication."	Çevik et al. (2025)
B2: Structured routines	10/10 (100%)	"Mentoring sessions followed a semester-based cycle with biweekly meetings, mid-term check-ins, and end-of-term reflection assignments."	Raymond et al. (2021)
C1: External pressures	7/10 (70%)	"The digital transformation mandate from headquarters created urgency for senior engineers to acquire new competencies from younger colleagues."	Gupta et al. (2024)
C2: Evaluation mechanisms	8/10 (80%)	"Pre-post surveys measured changes in digital literacy, reflective journals captured qualitative learning, and completion reports documented program outcomes."	Juris et al. (2022)
Outcomes			
O-L: Learning	10/10 (100%)	"Student mentors reported significant growth in communication skills, organizational awareness, and professional confidence through the mentoring experience."	Raymond et al. (2021)
O-A: Alignment	9/10 (90%)	"The program reduced informational distance between students and leadership; administrators reported gaining direct insight into student perspectives they would not encounter through formal reporting channels."	Gündüz & Akşit (2018)
O-I: Innovation	6/10 (60%)	"Reverse mentoring predicted employees' innovative behavior, with continuous learning serving as the mediating mechanism between mentoring participation and innovation outcomes."	Li et al. (2025)

Note: Quotations are paraphrased from original studies to illustrate coding decisions. Full coding dictionary with operational definitions is provided in Appendix B.

5. Discussion

5.1. Theoretical Contributions

The central theoretical contribution of this review lies in demonstrating how four previously separate theoretical perspectives operate as a sequential mechanism. This review validates and refines the conceptual framework proposed in the Literature Review while contributing specifically to understanding reverse

mentoring in higher education contexts. The thematic synthesis elaborated the initial broad categories of “governance, culture, and resources” into six operational factors: governance and culture (A1–A2), resources and structure (B1–B2), and external pressures and evaluation (C1–C2). This specification aligns with Ghosh and Reio’s (2013) argument that mentoring success depends on alignment across multiple organizational layers rather than isolated elements. Importantly, the factors showed differential salience across sectors—higher education demonstrated strength in governance, culture, structure, and evaluation, while corporate settings relied more heavily on external competitive pressures (Garg et al., 2021). The finding challenges the implicit assumption in early reverse mentoring literature that a single implementation model applies across contexts (Marcinkus Murphy, 2012), suggesting instead that effective implementation requires calibration to sector-specific profiles.

The outcome analysis revealed three distinct dimensions—learning, alignment, and innovation—that appear consistent with a developmental sequence. The proposed progression parallels Crossan et al.’s (1999) 4I framework, where individual intuiting and interpreting aggregate into group-level integration before achieving institutional change. However, our finding that higher education achieved strong learning and alignment but no documented innovation contrasts with corporate studies reporting innovation as a primary outcome (Li et al., 2025; Garg et al., 2021). This divergence may reflect measurement differences rather than actual outcome absence, or it may indicate that university programs have not yet achieved the institutionalization that organizational learning theory suggests requires sustained effort over extended timeframes (Argote & Miron-Spektor, 2011). This L→A→I sequence represents a theoretically grounded proposition rather than an empirically confirmed pathway. With only one longitudinal study (Li et al., 2025) in the corpus, cross-sectional evidence can demonstrate co-occurrence but not temporal ordering. Future research employing multi-wave designs is needed to test whether outcomes unfold sequentially as theorized.

The literature review presented Social Exchange (Chaudhuri & Ghosh, 2012), Social Learning (Bandura, 1977), Organizational Learning (Argyris & Schön, 1978), and Dynamic Capabilities (Teece, 2007) as separate theoretical lenses. The synthesis reveals how they operate as interconnected mechanisms: governance creates exchange conditions motivating participation (Chaudhuri & Ghosh, 2012); psychological safety enables observational learning (Bandura, 1977); training and evaluation convert individual insights into organizational knowledge (Argyris & Schön, 1978); and institutionalization generates renewal capabilities (Teece, 2007). This integration addresses Chaudhuri et al.’s (2022) critique that reverse mentoring research invokes theories selectively without articulating connections. The contribution lies not in identifying these perspectives but in tracing their sequential activation within a unified process model.

Beyond framework refinement, the review contributes higher education-specific understanding. The four adoption patterns identified—governance-oriented, training-centred, curriculum-embedded, and community-partnered—extend Marcinkus Murphy’s (2012) general framework to university contexts. The finding that students serve as “insider-outsiders” (Chun & Evans, 2021) resonates with Matthews and Dollinger’s (2023) argument that student voice mechanisms derive power from this dual positioning. However, unlike their focus on formal representation, reverse mentoring offers a complementary informal channel that may reach perspectives formal structures miss.

5.2. Cross-Sector Insights

The comparative analysis reveals complementary strengths across sectors, offering opportunities for mutual learning while also highlighting gaps in current understanding. Technology and manufacturing firms consistently documented innovation outcomes absent from higher education studies. Li et al. (2025) employed two-wave surveys with validated scales; Gupta et al. (2024) connected programs to digital transformation metrics. University studies captured learning through reflective journals without tracking whether insights translated into changed practices. This pattern echoes Chaudhuri et al.’s (2022) observation that reverse mentoring research privileges proximate indicators over distal organizational impacts. Universities might adapt corporate measurement approaches while retaining qualitative richness (Geller et al., 2025). More fundamentally, manufacturing firms’ strategic alignment of reverse mentoring—connecting it to transformation agendas rather than treating it as a standalone initiative—offers a sustainability model that prior higher education literature has not emphasized (Garg & Singh, 2020).

The SME study's "borrowed governance" model—using federation coordination to compensate for limited internal capacity (Schwartz et al., 2022)—suggests possibilities for resource-constrained institutions. This finding extends Joseph and Fazaeli's (2023) community-embedded mentoring concept to inter-organizational collaboration. Regional university consortia could share training and evaluation resources, democratizing access beyond well-resourced research universities.

Conversely, higher education's systematic evaluation culture produced stronger evidence bases than other sectors, where assessment remained informal. This advantage has gone largely unrecognized in cross-sector mentoring literature, which typically positions universities as learners from corporate practice (Ragins & Verbos, 2017). The academic calendar's structured rhythms similarly offer a transferable model for organizations lacking natural program cycles (O'Connor et al., 2025). Taken together, the findings suggest that knowledge transfer in reverse mentoring implementation should flow bidirectionally rather than assuming corporate superiority.

5.3. Interpreting the Innovation Gap in Higher Education

The most striking sectoral divergence concerns innovation outcomes: all six non-higher education studies reported innovation (100%), while none of the four higher education studies did (0%). This pattern warrants systematic examination through four complementary lenses.

1. *Measurement explanation.* Higher education studies primarily measured pedagogical and developmental outcomes—mentor skill growth, faculty digital competence, intergenerational understanding—rather than organizational innovation indicators. In contrast, Li et al. (2025) explicitly operationalized innovative behavior using validated scales. If higher education researchers had employed similar instruments, innovation outcomes might have been detected. Recent scholarship confirms that innovation measurement in universities requires sector-specific indicators beyond those developed for corporate contexts (Fernández et al., 2023).
2. *Timeframe explanation.* The higher education studies examined programs spanning single semesters, while corporate initiatives extended across multiple years. Institutional change in universities—characterized by shared governance and tenure systems—requires longer gestation periods than captured here. What appears as absence may reflect truncated observation windows (Alzahmi et al., 2025).
3. *Conceptual explanation.* In higher education, learning and alignment may constitute valued outcomes rather than precursors to innovation. When administrators gain unfiltered student perspectives (Gündüz & Akşit, 2018) or faculty develop digital competencies (Çevik et al., 2025), these directly serve institutional missions. The concept of "innovation" borrowed from corporate contexts may not translate directly to educational settings where relationship strengthening is a primary goal.
4. *Structural explanation.* Universities as professional bureaucracies differ fundamentally from corporate hierarchies (Mintzberg, 2024). Faculty autonomy and collegial governance mean insights from reverse mentoring must navigate multiple approval processes before becoming institutionalized. The pathway from learning to innovation may be longer and more indirect, requiring sustained effort across multiple program cycles (Kallio et al., 2020).

These explanations are not mutually exclusive. The finding should not be interpreted as evidence that reverse mentoring fails to produce innovation in higher education. Rather, it suggests future studies should employ innovation-specific measures, extend observation periods, adapt innovation concepts for educational contexts, and account for structural features that may delay but not prevent institutional change.

5.4. Practical Implications for Higher Education

Drawing on the findings, Box 1 synthesizes evidence-based guidance for university leaders. Each recommendation is tagged with an evidence strength indicator reflecting the consistency and directness of empirical support. University leaders should prioritize strongly supported elements while treating moderately supported recommendations as adaptable to local contexts.

Table 12. Implementing Reverse Mentoring in Higher Education - A Six-Step Guide

Step	Action	Evidence-Based Specifications	Evidence Strength
1. Establish Governance	Secure visible senior sponsorship; designate coordinating unit (TLC, Student Affairs, or governance office)	Presidential or provost-level endorsement documented in 100% of successful HE cases (Gündüz & Akşit, 2018; Çevik et al., 2025)	★★★ Strong
2. Design Pairing	Match based on complementary expertise; maintain 1:1 or 1:2 mentor-to-mentee ratios	Technology-based matching (Çevik et al., 2025) or strategic topic alignment (Raymond et al., 2021); larger ratios risk dilution	★★☆ Moderate
3. Train Mentors	Provide 4–8 hours pre-program orientation covering communication, boundaries, and institutional context	Structured training correlated with higher mentee satisfaction (Çevik et al., 2025); content should address power dynamics (Resch, 2023)	★★★ Strong
4. Structure Interaction	Align with semester (8–12 weeks); schedule biweekly 45–60-minute sessions; define discussion protocols	Academic calendar alignment reported across all HE studies; excessive flexibility reduces engagement	★★★ Strong
5. Evaluate Systematically	Deploy pre-post surveys + reflective journals; conduct mid-cycle check-ins; track completion rates	Multi-method evaluation standard in HE contexts (100% of studies); single-method approaches miss process insights	★★★ Strong
6. Institutionalize	Document procedures; train successive cohorts; connect to strategic priorities; report to governance bodies	Strategic alignment with institutional goals predicts sustainability (Gupta et al., 2024); isolated pilots rarely persist	★★☆ Moderate

Evidence strength key:

- ★★★ Strong = Consistent finding across 3+ studies with direct empirical support
- ★★☆ Moderate = Supported by 1–2 studies or inferred from cross-sector evidence
- ★★☆ Emerging = Theoretically grounded but limited empirical validation

5.5. Policy Implications

Beyond institutional implementation, the findings carry implications for higher education policy. First, quality assurance frameworks increasingly emphasize student engagement; reverse mentoring offers a mechanism that moves beyond consultation toward genuine partnership in governance (Matthews & Dollinger, 2023). Accreditation bodies and ministries might recognize structured reverse mentoring as evidence of meaningful student voice, incentivizing adoption. Second, the consortium model documented in SME contexts (Schwartz et al., 2022) suggests policy opportunities for regional coordination. Higher education authorities could facilitate shared training resources and evaluation instruments across institutions, reducing barriers for smaller universities lacking teaching and learning centers. Third, the innovation gap identified in this review points toward a research policy priority: funding agencies should encourage longitudinal studies tracking whether reverse mentoring contributes to institutional innovation over multi-year timeframes. Such investment would address the methodological limitations constraining current understanding while generating actionable evidence for practice (Palma et al., 2023).

5.6. Limitations and Future Research Directions

Several constraints shape interpretation of these findings and suggest research priorities. The review included only ten studies, with four higher education cases concentrated in Turkey and the United States. This geographic limitation is notable given evidence that mentoring practices vary significantly across cultural contexts (Ragins & Kram, 2007). The Asia-Pacific region, despite its large higher education sector, contributed no studies. Future research should examine whether adoption patterns and outcome sequences hold across different governance traditions and institutional cultures, or whether alternative configurations emerge in Confucian, South Asian, or other educational contexts.

The open-access filter applied during screening introduced additional limitations. As documented in Table 4, approximately 12 relevant empirical studies published in subscription journals could not be accessed for full-text review. These include quantitative research from India examining reverse mentoring's effects on innovation and turnover (Garg et al., 2021; Upadhyay & Singh, 2024), qualitative studies from Poland on intergenerational relationships (Gadomska-Lila, 2020), and a higher education case study from the UK on student-leadership mentoring (O'Connor et al., 2025). The paywalled studies are concentrated in European management journals and quantitative research designs, suggesting our findings may under-represent these perspectives. The cross-sector comparison particularly may not reflect typical corporate practice, as premier management journals such as *Human Resource Management* and *Career Development International* rarely offer immediate open-access options.

The methodological profile—seven qualitative cases, two cross-sectional surveys, one longitudinal study—limits causal inference. The learning–alignment–innovation sequence represents a theoretical proposition supported by cross-sectional patterns but not yet validated through longitudinal evidence. Only Li et al. (2025) employed a multi-wave design; the remaining studies cannot establish whether learning precedes alignment or whether alignment precedes innovation. Alternative temporal orderings—such as innovation attempts driving subsequent learning, or simultaneous rather than sequential emergence—cannot be ruled out with current evidence.

6. Conclusion

This systematic review synthesized ten empirical studies on reverse mentoring to examine how this practice is adopted, implemented, and linked to outcomes in higher education. By positioning four university-based studies as the primary analytical focus and six studies from other sectors as comparative reference, the review offers both a focused understanding of higher education contexts and cross-sector insights that enrich practical guidance.

The analysis reveals that universities approach reverse mentoring in ways shaped by their institutional character. Four distinct adoption patterns emerged—governance-oriented, training-centred, curriculum-embedded, and community-partnered—yet all shared reliance on existing infrastructure, academic calendars, and evaluation traditions that distinguish higher education from corporate settings. Students occupy a particularly valuable position as institutional insiders who nonetheless bring outsider perspectives, creating a tension that employee-to-employee mentoring cannot replicate.

Six contextual enablers—governance, culture, resources, structure, external pressures, and evaluation—together shape whether reverse mentoring takes root and endures, with universities demonstrating strength in most factors but relative weakness in external pressure drivers compared to corporate sectors where competitive imperatives create urgency. Outcomes followed a developmental pattern: learning appeared consistently across all higher education studies, alignment built upon learning foundations in most cases, while innovation outcomes remained largely undocumented in university research despite their prominence in corporate contexts—a gap that may reflect measurement differences, shorter study timeframes, or genuine distinctions in how universities pursue organizational change. Cross-sector comparison proved illuminating in both directions: technology firms' quantitative outcome tracking, manufacturing organizations' strategic alignment with transformation agendas, and SME consortium resource-sharing each offer models universities might adapt, while higher education's systematic evaluation culture and structured temporal rhythms represent transferable assets that corporate implementations often lack.

For university leaders, the findings suggest treating reverse mentoring as a strategic mechanism for institutional learning and participatory governance rather than a peripheral experiment. Teaching and learning centres emerge as natural governance anchors for such initiatives, offering coordination capacity, training expertise, and evaluation infrastructure. Regional consortia may enable resource-constrained institutions to participate through shared investment, and student shadow boards may strengthen governance feedback beyond traditional representative structures. The review's limitations—a small sample concentrated in Turkey and the United States, predominance of qualitative designs limiting causal inference, and absence of longitudinal tracking—point toward priorities for future research. Empirical studies from underrepresented regions, particularly the Asia-Pacific, would test whether identified patterns generalize across institutional cultures; longitudinal mixed-methods designs tracking participants over multiple academic years would strengthen

confidence in the learning–alignment–innovation pathway. Whether and how reverse mentoring translates into broader institutional transformation remains a question warranting sustained attention from both researchers and practitioners.

7. Ethics Statement

This systematic review synthesized published empirical studies and did not involve primary data collection from human subjects. Therefore, no ethical approval was required.

8. Data Availability Statement

The quality appraisal scores are presented in Appendix A. The coding dictionary with operational definitions, inclusion criteria, and decision rules is provided in Appendix B. The complete data extraction matrix showing all included studies with full citations, country, sector, research design, sample, main findings, enabler codes (A1–C2), outcome codes (O-L/O-A/O-I), and CCAT scores is available as Supplementary Material. The NVivo codebook export is available from the corresponding author upon reasonable request.

9. Declaration

The authors acknowledge the use of artificial intelligence tools (ChatGPT, OpenAI) to assist with language editing and clarity improvement during the preparation of this manuscript. The AI tool was not used for data analysis, interpretation of results, or generation of substantive scientific content. All final decisions regarding the content remain the responsibility of the authors.

References

- Alzahmi, R. A., Syed, R. T., Singh, D., Arshi, T. A., & Kutty, S. V. (2025). Organizational change in higher education institutions: thematic mapping of the literature and future research agenda. *Humanities and Social Sciences Communications*, 12(1), 1-15. <https://doi.org/10.1057/s41599-025-05650-w>
- Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, 22(5), 1123–1137. <https://doi.org/10.1287/orsc.1100.0621>
- Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Çevik, Y. D., Şen, U. U., & Avcı, D. (2025). Evaluating the Impact of an Online Course for Technology Mentors in Reverse Mentoring. *Journal of Learning and Teaching in Digital Age*, 10(2), 244-259. <https://doi.org/10.53850/joltida.1627427>
- Chaudhuri, S., & Ghosh, R. (2012). Reverse mentoring: A social exchange tool for keeping the boomers engaged and millennials committed. *Human Resource Development Review*, 11(1), 55–76. <https://doi.org/10.1177/1534484311417562>
- Chaudhuri, S., Park, S., & Johnson, K. R. (2022). Engagement, inclusion, knowledge sharing, and talent development: Is reverse mentoring a panacea to all? Findings from literature review. *European Journal of Training and Development*, 46(5/6), 518–539. <https://doi.org/10.1108/EJTD-01-2021-0005>
- Chen, Y.-C. (2013). Effect of reverse mentoring on traditional mentoring functions. *Leadership and Management in Engineering*, 13(3), 199–208. [https://doi.org/10.1061/\(ASCE\)LM.1943-5630.0000227](https://doi.org/10.1061/(ASCE)LM.1943-5630.0000227)
- Chun, E., & Evans, A. (2021). Leveraging multigenerational workforce strategies in higher education. Routledge. <https://doi.org/10.4324/9781003150381>
- Cismaru, L., & Iunius, R. (2020). Bridging the generational gap in the hospitality industry: Reverse mentoring—An innovative talent management practice for present and future generations of employees. *Sustainability*, 12(1), 263. <https://doi.org/10.3390/su12010263>
- Crossan, M. M., Lane, H. W., & White, R. E. (1999). An organizational learning framework: From intuition to institution. *Academy of Management Review*, 24(3), 522–537. <https://doi.org/10.5465/amr.1999.2202135>
- Crowe, M., & Sheppard, L. (2011). A review of critical appraisal tools show they lack rigor: Alternative tool for evaluating health research. *BMC Medical Research Methodology*, 11, 229. <https://doi.org/10.1016/j.jclinepi.2010.02.008>
- Dale, C. A. (2025). *The phenomenology of intelligences: Connecting personal flexibility with corporate agility in complexity* (Doctoral dissertation, Middlesex University). <https://repository.mdx.ac.uk/item/2x03v7>
- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L., Riley, R., & Sutton, A. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology*, 6(35). <https://doi.org/10.1186/1471-2288-6-35>
- Fernández, A., Gómez, B., Binjaku, K., & Meço, M. (2023). Digital transformation initiatives in higher education institutions: A multivocal literature review. *Education and Information Technologies*, 28, 12351–12382. <https://doi.org/10.1007/s10639-022-11544-0>

- Gadomska-Lila, K. (2020). Effectiveness of reverse mentoring in creating intergenerational relationships. *Journal of Organizational Change Management*, 33(6), 1135–1153. <https://doi.org/10.1108/JOCM-10-2019-0326>
- Garg, N., Murphy, W., & Singh, P. (2021). Reverse mentoring, job crafting and work-outcomes: the mediating role of work engagement. *Career Development International*, 26(2), 290–308. <https://doi.org/10.1108/cdi-09-2020-0233>
- Garg, N., & Singh, P. (2020). Reverse mentoring: a review of extant literature and recent trends. *Development and Learning in Organizations: An International Journal*, 34(5), 5–8. <https://doi.org/10.1108/dlo-05-2019-0103>
- Garvey, B., Stokes, P., & Megginson, D. (2017). *Coaching and mentoring: Theory and practice* (3rd ed.). SAGE Publications.
- Geller, H. E., Golding, P., Perez, A., & Golding, D. E. (2025). Exploring integrated peer and reverse mentoring in engineering education: A work in progress. *2025 ASEE Annual Conference*.
- Ghosh, R., & Reio, T. G. (2013). Career benefits associated with mentoring for mentors: A meta-analysis. *Journal of Vocational Behavior*, 83(1), 106–116. <https://doi.org/10.1016/j.jvb.2013.03.011>
- Gong, Z., Wu, J. & Li, M. (2024). Reverse Mentoring in an Aging Workplace: Research Review and Future Prospects. *Miaomiao, Reverse Mentoring in an Aging Workplace: Research Review and Future Prospects*. <https://doi.org/10.2139/ssrn.4715659>
- Gordon, P. (2025). Reverse Mentoring: Transforming Learning at Individual, Team, and Organizational Levels to Build Belonging in the Workplace. *International Journal of Human Resource Development: Practice, Policy and Research*, 9(1), 53-60. <https://doi.org/10.2478/ijhrd-2025-0006>
- Gündüz, Ş., & Akşit, B. (2018). Student-president reverse mentoring at universities: Maltepe University case. *Journal of Higher Education*, 8(3), 346–356. <https://doi.org/10.2399/yod.18.019>
- Gupta, S., Modgil, S., Bhushan, B., Kamble, S., & Mishra, J. (2024). Digital innovation and transformation capabilities in a large company. *Expert Systems*, 41(1), e13452. <https://doi.org/10.1111/exsy.13452>
- Harvey, C. P., & Allard, M. J. (2020). *Understanding and managing diversity: Readings, cases, and exercises* (7th ed.). Pearson.
- Joseph, D., & Fazaeli, F. (2023). Community-based mentoring: Exploring non-traditional models. *Mentoring & Tutoring: Partnership in Learning*, 31(2), 156–174. <https://doi.org/10.1080/13611267.2023.2189432>
- Juris, J. J., Bouldin, E. D., Uva, K., Cardwell, C. D., Schulhoff, A., & Hiegl, N. (2022). Virtual intergenerational reverse-mentoring program reduces loneliness among older adults: Results from a pilot evaluation. *International Journal of Environmental Research and Public Health*, 19(12), 7121. <https://doi.org/10.3390/ijerph19127121>
- Kallio, T. J., Kallio, K.-M., & Blomberg, A. (2020). From professional bureaucracy to competitive bureaucracy—Redefining universities' organization principles, performance measurement criteria, and reason for being. *Qualitative Research in Accounting & Management*, 17(1), 82–108. <https://doi.org/10.1108/gram-10-2019-0111>
- Li, M., Guo, J., Zou, C., & Yin, J. (2025). The impact of reverse mentoring on employees' innovative behavior: Evidence from Chinese technology enterprises. *Sustainability*, 17(1), 6. <https://doi.org/10.3390/su17010006>
- Lužar, M., Gorenc Zoran, A., Markič, M., & Bukovec, B. (2023). Intergenerational differences and knowledge transfer among Slovenian engineers. *Organizacija*, 56(1), 65–82. <https://doi.org/10.2478/orga-2023-0005>

- Marcinkus Murphy, W. (2012). Reverse mentoring at work: Fostering cross-generational learning and developing millennial leaders. *Human Resource Management*, 51(4), 549–573. <https://doi.org/10.1002/hrm.21489>
- Matthews, K. E., & Dollinger, M. (2023). Student voice in higher education: The importance of distinguishing student representation and student partnership. *Higher Education*, 85, 550–570. <https://doi.org/10.1007/s10734-022-00851-7>
- Mintzberg, H. (2024). Four forms that fit most organizations. *California Management Review*, 66(2), 5–28. <https://doi.org/10.1177/00081256231214816>
- O'Connor, R., Barraclough, L., Gleadall, S., & Walker, L. (2025). Institutional reverse mentoring: Bridging the student/leadership gap. *British Educational Research Journal*, 51(1), 344–368. <https://doi.org/10.1002/berj.4078>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372. <https://doi.org/10.1136/bmj.n71>
- Palma, A., Mourato, J., Vinagre, J., Almeida, F., & Pessoa, A. M. (2023). Student participation: Issues for the governance of higher education. *Societies*, 13(5), 115. <https://doi.org/10.3390/soc13050115>
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. London: SAGE Publications.
- Ragins, B. R., & Kram, K. E. (2007). *The handbook of mentoring at work: Theory, research, and practice*. Sage Publications.
- Ragins, B. R., & Verbos, A. K. (2017). Positive relationships in action: Relational mentoring and mentoring schemas in the workplace. In *Exploring positive relationships at work* (pp. 91–116). Psychology Press.
- Raymond, M. A., Siemens, J., & Thyroff, A. (2021). Students, please teach us! Implementing student-employee reverse mentoring to increase career readiness. *Marketing Education Review*, 31(2), 87–92. <https://doi.org/10.1080/10528008.2021.1907593>
- Rutti, R. M., Helms, M. M., & Rose, L. C. (2013). Changing the lens: Viewing the mentoring relationship as relational structures in a social exchange framework. *Leadership & Organization Development Journal*, 34(5), 446–468. <https://doi.org/10.1108/LODJ-11-0097>
- Saidi, Z. A. & Abd Aziz, M. F. (2025). Empowering student leadership in higher education: A literature review on key competencies for effective leadership. *Journal of Institutional Research South East Asia*, 23(1), 376–399.
- Saidi, Z.A., Asri, A.S., & Abd Aziz, M.F. (2025). Exploring the relationship between participation, needs, readiness, and planning in university student volunteering programs. *Journal of Institutional Research South East Asia*, 23(1), 401–417.
- Syed, R. T., & Spicer, D. (2025). Organizational change in higher education institutions: Thematic mapping of the literature and future research agenda. *Humanities and Social Sciences Communications*, 12, 1087. <https://doi.org/10.1038/s41599-025-05650-w>
- Schwartz, D., Bar-El, R., & Bentolila, D. J. (2022). Adapting reverse mentoring strategy to SMEs: A new pilot model implemented in Brazil. *Sustainability*, 14(15), 9515. <https://doi.org/10.3390/su14159515>
- Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319–1350. <https://doi.org/10.1002/smj.640>

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, Article 45. <https://doi.org/10.1186/1471-2288-8-45>

Yurtseven Avici, Z., & Eren, E. (2023). Intergenerational Interdisciplinary Reverse Mentoring: School-University Collaboration. *Hacettepe University Journal of Education*. LOCKSS. <https://doi.org/10.16986/huje.2023.497>

Zauchner-Studnicka, S. A. (2017). A model for reverse-mentoring in education. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 11(3), 546-553.

APPENDIX

Appendix A: CCAT Quality Appraisal Scores

Study	Sector	Prel.	Intro.	Design	Sampling	Data	Ethics	Results	Disc.	Total (/40)
Gündüz & Akşit (2018)	HE	4	4	4	3	4	3	4	4	30
Çevik et al. (2025)	HE	5	4	4	4	5	4	4	4	34
Raymond et al. (2021)	HE	4	4	3	3	4	3	4	4	29
Juris et al. (2022)	HE	4	4	4	4	4	4	4	4	32
Chen (2013)	Tech	4	4	4	3	4	3	4	4	30
Li et al. (2025)	Tech	5	5	5	5	5	4	5	5	39
Gupta et al. (2024)	Mfg	4	5	4	4	4	4	5	4	34
Lužar et al. (2023)	Mfg	4	4	4	4	4	3	4	4	31
Cismaru & Iunius (2020)	Hosp	4	4	4	3	4	3	4	4	30
Schwartz et al. (2022)	SME	4	4	4	4	4	3	4	4	31

Note: HE (Higher Education), Tech (Technology Enterprises), Mfg (Large Manufacturing), Hosp (Hospitality & Services), SME (Small Medium-sized Enterprises)

Appendix B: Coding Dictionary

This appendix provides the operational definitions and coding rules used in the thematic synthesis. The coding framework comprises six contextual enablers (Table B1) and three outcome themes (Table B2), with decision rules specified in Table B3.

Table B1. Contextual Enablers – Operational Definitions and Coding Rules

Code	Definition	Inclusion Criteria	Example Evidence
A1-Governance	Formal leadership support, institutional arrangements, and authority allocation for RM programs	Explicit mention of senior sponsorship, policy backing, designated coordinating unit, or resource commitment from leadership	"The president directly endorsed the program"; "TLC serves as the coordinating unit"
A2-Culture	Organizational climate characterized by openness, psychological safety, and intergenerational respect	References to trust-building, openness to feedback from junior members, reduced hierarchy, or deliberate cultivation of safe dialogue	"Faculty felt comfortable receiving guidance from students"; "Open communication was encouraged"
B1-Resources	Dedicated investments including training programs, technology platforms, time allocation, and funding	Mention of budget allocation, training modules, learning management systems, or protected time for participation	"A 5-week online training course was developed"; "Funding was provided by the federation"
B2-Structure	Operational design elements including pairing mechanisms, interaction protocols, and program timelines	Description of matching criteria, meeting schedules, program duration, or structured interaction frameworks	"Mentors met with faculty biweekly for 8 weeks"; "Pairs were matched based on technology needs"

C1-External	Environmental pressures including policy mandates, competitive forces, and social mission imperatives	References to external mandates, industry competition, accreditation requirements, or community partnership obligations	"The program addressed regional aging policy goals"; "Digital transformation pressure drove adoption"
C2-Evaluation	Assessment mechanisms including measurement tools, feedback systems, and continuous improvement processes	Mention of surveys, interviews, reflection journals, outcome tracking, or program refinement based on data	"Pre-post questionnaires measured skill development"; "Feedback was used for program improvement"

Table B2. Outcome Themes – Operational Definitions and Coding Rules

Code	Definition	Strength Criteria	Example Evidence
O-L Learning	Individual and group-level knowledge acquisition, skill development, and perspective transformation	Strong: Multiple learning types (cognitive, behavioral, attitudinal) documented Moderate: Single learning type reported Weak: Learning implied but not explicitly measured	"Mentors reported improved communication skills"; "Faculty gained digital literacy"
O-A Alignment	Enhanced cross-level communication, improved coordination, and strengthened intergenerational relationships	Strong: Documented behavioral changes in communication patterns Moderate: Attitudinal improvements reported Weak: Relationship benefits implied only	"Information distance between students and leadership decreased"; "Cross-generational dialogue became routine"
O-I Innovation	Organizational-level process improvements, institutional changes, and creative outputs attributed to RM	Strong: Documented organizational or process changes Moderate: Individual innovative behaviors reported Weak: Innovation potential noted but not measured	"New teaching methods were adopted"; "Governance procedures were modified based on student input"

Table B3. Coding Decision Rules

Code	Decision Rule
Y (Yes)	Explicit evidence present in the study text; factor or outcome is directly described or measured
Partial	Indirect evidence or implied presence; factor mentioned but not central to study design or findings
N (No)	Not mentioned in the study or explicitly reported as absent

Two Decades of TPACK Research in Higher Education: A Global Bibliometric Review from 2006 To 2025

ABSTRACT

Authors: Noor Ibtisam Abdul Karim & Mohd Hazwan Mohd Puad*

Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:

hazwan@upm.edu.my

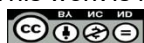
Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

The Technological Pedagogical and Content Knowledge (TPACK) has become a key model for understanding technology integration in higher education, particularly as universities undergo rapid digital transformation. This study conducts a bibliometric analysis of TPACK-related publications in higher education from 2006 to 2025, covering two decades of research since the framework's introduction. Data was retrieved from the Scopus database and analyzed using descriptive bibliometric techniques and science mapping with VOSviewer. The analysis examines publication trends, geographical distribution, subject areas, citation patterns, and keyword co-occurrence networks. Findings indicate a continuous growth in TPACK research, with a notable increase after 2018 and during the post-pandemic period. The United States, China, and Indonesia emerge as leading contributors, while Social Sciences and Computer Science dominate the subject areas. Highly cited studies mainly focus on technology acceptance, digital competence, and teacher readiness. Keyword co-occurrence analysis reveals strong connections between TPACK, higher education, digital competence, online learning, and technology integration. Overall, the findings suggest that TPACK has evolved into a widely applied and interdisciplinary framework in higher education, although gaps remain in its contextual application across disciplines and regions. This study provides insights to inform future research and support the continued development of digital teaching practices in higher education.

Keywords: Bibliometric analysis, digital pedagogy, higher education, TPACK.

1. Introduction

Over the past two decades, the rapid expansion of digital learning ecosystems has fundamentally modified the higher education landscape. Globally, universities are increasingly expected to integrate advanced digital technologies (Aditya et al., 2020; Basilotta-Gómez-Pablos et al., 2022) into teaching, learning, and assessment practices as part of broader digital transformation agendas. This shift has urged the need to understand how higher education instructors develop, apply, and sustain the knowledge required to teach effectively with technology. Beyond technological knowledge, recent higher education research has emphasized that effective teaching also requires a broader set of professional competencies, including communication, interpersonal skills, critical thinking, emotional intelligence, and digital proficiency, which collectively shape educators' ability to integrate technology meaningfully into teaching practices (Saidi & Abd Aziz, 2025). Driven by ongoing technological advancements, increased accessibility to digital tools, and national-level digital education policies, educators are now required to navigate the complex interplay between pedagogy, content, and technology to create meaningful learning experiences for diverse learners (Sanchez, 2020). As traditional pedagogical models are transformed into digitally enhanced contexts, theoretical frameworks that explain the complexity of instructional decision-making have become essential.

The Technological Pedagogical and Content Knowledge (TPACK), introduced by Mishra and Koehler in 2006, has emerged as one of the most influential and widely used models for conceptualizing teachers' technology integration competence. TPACK describes the interplay between three core forms of teacher knowledge (i.e. technology, pedagogy, and content) and articulates how their intersections shape technology enhanced teaching (Mishra & Koehler, 2006). Over time, TPACK has evolved from a conceptual model to a widely adopted research framework used across teacher education, professional development, curriculum design, and digital competence assessment. Many educational systems now embed TPACK competencies into teacher training programs, national ICT in education standards, and digital transformation initiatives. This increasing institutional and policy attention underscores TPACK's central role in guiding digital competence development for educators.

Beyond TPACK, several frameworks such as the European DigCompEdu, ISTE Standards, and the Professional Digital Competence (PDC) model have expanded the conceptualization of digital teaching competence by incorporating elements such as digital ethics, data literacy, accessibility, and institutional support. The models also show the interconnected nature of technology, pedagogy, and subject content in defining digital readiness. These frameworks highlight that effective technology integration requires not only knowledge of pedagogy and content, but also broader competencies related to digital citizenship (Redecker & Punie, 2017). While TPACK offers theoretical understanding of the knowledge intersections, these frameworks provide more towards procedural steps and levels of technology used in teaching and learning. In this context, TPACK continues to function as a foundational model, yet its relationship with emerging digital competence frameworks remains theoretically underexplored. The dominance of TPACK was perceived primarily as cognitive framework, while the occurrence of other complementary frameworks in the literature has conceptually shifted the operationalization of TPACK that requires more holistic towards digital competence as a part of educators' professional profile (Lindfors et al., 2021).

Despite the extensive use of the Technological Pedagogical Content Knowledge (TPACK) in educational research, existing systematic and bibliometric reviews have mainly focused on school education, pre-service teachers, or specific subject domains rather than higher education contexts. For instance, Shabira et al. (2024) reviews TPACK trends specifically in elementary education, while Zou et al. (2022) and Suprpto et al. (2021) did not specify any educational level in their review studies. In fact, these reviews primarily examine conceptual development, measurement approaches, and general research trends in TPACK, often within limited time spans or pedagogical settings that differ substantially from university teaching environments.

The application of TPACK in higher education has been documented as a framework for assessing educators' readiness to integrate technology into their teaching practices. Beyond evaluating individual readiness, TPACK has also been used to examine technology-oriented professional competence, assess the effectiveness of professional development programmes, and explore how digital tools support discipline-specific pedagogical approaches (Cabero & Barroso, 2016). At a broader level, the framework has been applied to understand shifts in instructional practices driven by institutional digital transformation initiatives within higher education. This trajectory was further accelerated during the COVID-19 pandemic, as universities rapidly transitioned to online, blended, and hybrid learning modes. In response, researchers increasingly employed TPACK to explain how educators adapted their teaching strategies, redesigned learning activities, and navigated emerging digital pedagogical challenges (Ali, 2020; Kallou & Kikilia, 2021).

The ecosystem of higher education is characterised by disciplinary diversity, instructional autonomy, and the research-teaching nexus, all of which influence how technology is integrated into teaching practices. Furthermore, unlike school teachers, university educators often enter academia with strong disciplinary expertise but limited formal pedagogical training (Mayasari et al., 2024). Due to these unique characteristics of higher educational institution settings, the application of TPACK might vary in some circumstances. Besides, the use of TPACK in higher education related studies have expanded substantially, incorporating diverse fields such as STEM, humanities, social sciences, vocational education, and professional training (Hsu et al., 2015). More recently, the emergence of generative artificial intelligence (AI), immersive technologies such as VR/AR, and the growing use of learning analytics have transformed scholarly interest in how TPACK can be reconceptualised or extended to remain relevant within rapidly evolving digital networks. The ongoing expansion reflects both the adaptability of the TPACK and the increasing complexity of digital teaching practices in contemporary higher education. Another failing factor to the progress of digital pedagogy in higher education is the COVID-19 pandemic, revealing the need to change and adapt to new digital things and environments (Chen & Hsu, 2022). These contextual differences pointed that TPACK may operate differently in the pedagogical and professional characteristics in higher education compared to structured school environments. As a result, the application of TPACK in higher education pedagogy becomes more complex and context-dependent, requiring a clearer understanding of how the framework has evolved within tertiary education research. Understanding how TPACK scholarship has evolved across higher education is therefore essential for positioning the model within the broader landscape of digital pedagogy.

The significant gap of a focused bibliometric synthesis on TPACK in higher education has resulted in a fragmented understanding of publication patterns, influential contributors, thematic development, and intellectual structures within this field. Addressing this gap, the present study conducts a comprehensive higher-education-focused bibliometric review of TPACK research published between 2006 and 2025 as well as using co-occurrence network analyses. Specifically, the study aims to map the evolution of TPACK research in higher education, identify major research clusters and influential contributions, and examine emerging directions that shape TPACK scholarship at the tertiary level.

To guide the interpretation of bibliometric findings, this study adopts an explicit conceptual framework that links bibliometric indicators to theories of knowledge development and scientific evolution. The growth of publications over time can be understood as reflecting the progression of a research field from its emergence to consolidation and expansion. In this context, increasing publication output indicates the normalization of TPACK as a dominant framework within higher education research, while shifts in thematic focus may signal evolving research paradigms in response to technological and pedagogical change.

In addition, citation analysis is interpreted through the lens of citation network, which views citations as indicators of knowledge diffusion, intellectual influence, and scholarly recognition within a field.

Highly cited publications represent foundational or influential works that shape subsequent research directions, while normalized citation metrics (e.g., citations per year) provide insights into emerging influential studies.

Furthermore, keyword co-occurrence analysis is grounded in knowledge domain evolution models, where co-occurring terms reflect the conceptual structure and thematic organization of a research field. Clusters of keywords represent interconnected research themes, while link strength indicates the degree of conceptual integration across topics. Together, these bibliometric indicators provide a multi-dimensional understanding of how TPACK research in higher education has developed, diversified, and interconnected over time. This framework enables the study to move beyond descriptive mapping toward a more analytical interpretation of the field's intellectual structure and evolution. By integrating these perspectives, the study positions bibliometric analysis not merely as a descriptive tool, but as an analytical approach for examining the dynamics of knowledge production and transformation within higher education research.

To date, bibliometric research on digital pedagogy has been studied in various angles of higher education context. Santoveña-Casal & Lopez (2024) focused on digitalisation, education quality, and university education, and some investigate smart learning (Rad et al., 2025), digital transformation in developing countries (Omar & Abdullahi, 2024) or digital literacy (Barrios-del-Ángel et al., 2026).

This paper eventually focuses on the following specific research objectives:

1. examine the publication output and growth trends of TPACK research in higher education between 2006 and 2025;
2. identify the most influential countries, journals, and authors contributing to TPACK scholarship in higher education;
3. analyse the subject areas characterize TPACK research in higher education;
4. determine the most highly cited publications shaping the development of TPACK research in higher education; and
5. explore the intellectual structure among key TPACK-related concepts and higher education using keyword co-occurrence analysis

To achieve these objectives, the following research questions were considered:

1. RQ1: How has TPACK research in higher education evolved in terms of publication output and growth trends between 2006 and 2025?
2. RQ2: Which countries, journals, and authors have contributed most significantly to TPACK scholarship in higher education?
3. RQ3: What subject areas characterize TPACK research in higher education?
4. RQ4: Which documents have exerted the greatest scholarly influence within the field based on citation patterns?
5. RQ5: How are key concepts related to TPACK and higher education interconnected within the related constructs of literature?

Through performance analysis, co-authorship networks, keyword co-occurrence mapping, and citation formations, this review provides a comprehensive overview of the field's growth, intellectual foundations, and shifting research primacies. By synthesizing approximately two decades of scholarship, the study intends to highlight progress, reveal persistent gaps, and outline opportunities for enhancing TPACK research in higher education as universities continue to deepen their digital transformation initiatives. This technique will be able to show what's been done, what's still missing, and how TPACK research can be advanced in tertiary education. Hence, this study offers timely and critical insights for researchers, educators, policymakers, and institutions seeking to understand how

TPACK has developed and how it can continue to inform digital teaching competence in an era of rapid educational transformation.

2. Methodology

This study employed a bibliometric research design to systematically map the scientific structure, output patterns, and conceptual development of TPACK research, specifically in higher education. Bibliometric methods provide quantitative insights into publication growth, citation patterns, research productivity, and global knowledge structures across identified body of literature. The analysis combined descriptive performance indicators (e.g., publication output, citations, subject areas) with science mapping techniques using VOSviewer. This study analysed bibliographic metadata retrieved from the existing Scopus database and did not involve human participants, personal data, or institutional interventions. Table 1 summarize the correspondence between research questions, data source and bibliometric techniques used in this study. The bibliometric data were retrieved from Scopus database in CSV format including the title, authors, publishers, keywords and others.

Table 1. Correspondence between Research Questions, Data Source, Bibliometric Techniques and Output in Results

Research Question	Data Source	Bibliometric Technique	Output in Results	Theoretical Interpretation
RQ1: Publication Trends	Scopus metadata (year)	Descriptive analysis and trend analysis	Annual publication trend	Evolution of TPACK as a research domain
RQ2: Contributors (Countries, Authors, Journals)	Author affiliations, source titles	Performance analysis	Top contributors	Knowledge production and global diffusion
RQ3: Subject Areas	Scopus subject categories	Descriptive classification	Subject distribution	Disciplinary spread of TPACK
RQ4: Influential Documents	Citation counts	Citation analysis	Top-cited publications	Intellectual influence and knowledge base
RQ5: Conceptual Relationships	Author keywords	Co-occurrence analysis (VOSviewer)	Keyword network map	Intellectual structure and thematic clusters

2.1 Search Strategy

The search strategy was performed to get the aimed literature, conducted using the Scopus database (Elsevier) via the advanced search interface on 15 November 2025 (10:20 GMT+8). The strategic search string was applied to the TITLE-ABS-KEY fields using the following query: TITLE-ABS-KEY ("TPACK" OR "technological pedagogical content knowledge" OR "technological pedagogical and content knowledge") AND TITLE-ABS-KEY ("higher education" OR "university" OR "tertiary education" OR "postsecondary" OR "college instructor*" OR "university lecturer*") AND PUBYEAR > 2005 AND PUBYEAR < 2026.

This search resulted 976 records. All retrieved records were exported from Scopus into CSV format to ensure compatibility with descriptive analysis and bibliometric mapping. The exported metadata included: document title, authors, year of publication, author affiliations, source title, abstract, author keywords, subject areas, and bibliographic references. The exported files were saved as TPACK_HE_Scopus_2006–2025_15nov2025.csv. Two duplicate records were identified using Scopus document identifiers (EID and DOI) and removed from the list. Further, screening, eligibility and inclusion steps were justified and in Table 2, explaining the exclusion of non-relevant records. The DOI list for each included document is supplied as supplementary material. The screening process was conducted by a pair researcher, independently using predefined inclusion and exclusion criteria, ensures that bibliometric metadata selection is objective, reproducible, and accurate. A subset of records (10%)

was rechecked after initial screening to verify accuracy. Then, any ambiguous cases were reviewed and discussed carefully to resolve any disagreement and confirm their relevance to the study contexts.

Table 2. Selection Process

Stage	Number of Records	Description
Records identified through Scopus search	976	Initial retrieval using search query
Records after duplicates removed	974	2 duplicate records removed
Records screened (title & abstract)	974	Eligibility screening conducted
Records excluded (based on exclusion criteria): Non-English publications Not relevant to higher education (e.g., school-level or non-educational domain)	124 151	Language restriction applied Context-based exclusion
Final included studies	699	Studies retained for analysis

2.2 Eligibility Criteria for Higher Education and English Language

To ensure consistency in data selection for this study, “higher education” was defined operationally by combining the keyword-based flitting and manual screening. Specifically, the search query included terms such as “higher education,” “university,” “tertiary education,” “postsecondary,” “college instructor,” and “university lecturer” within the title, abstract, and keyword fields in Scopus. These terms were used to capture studies explicitly situated within tertiary education contexts. However, due to the teaching environment context, some secondary level documents appeared. Therefore, manual screening has been conducted to all retrieved records to confirm their relevance with the anticipated scope. Studies focusing exclusively on primary or secondary school contexts, teacher education programmes without clear links to university-level instruction, or non-formal education environments were excluded. For ambiguous cases (e.g., studies involving both school and university participants), inclusion was determined based on whether higher education constituted the primary research context. Further, no restriction based on Scopus subject area classification was applied, as TPACK research spans multiple disciplines. This combined approach ensured both systematic retrieval and contextual relevance of the final dataset.

The analysis was limited to documents indexed in the Scopus database and written in English. The English-language filtration was applied to ensure consistency in keyword clarification and analytical accuracy. Stability and solidity of Scopus metadata with English language focus stands the bibliometric review technique to be more realistic and accurate. Besides, Scopus was selected due to its broad multidisciplinary coverage and standardized indexing of citation metadata, which supports reliable bibliometric analysis. Scopus comprehensive coverage of peer-reviewed journals and its extensive indexing of educational technology and higher education research also supplemented the objectives of the study.

2.3 Vosviewer settings

The keyword co-occurrence analysis was conducted using VOSviewer (version 1.6.20). Full counting was applied, whereby each co-occurrence of keywords was counted equally across publications. This approach was selected as it allows for a more direct representation of the frequency and strength of relationships between keywords, which is appropriate for identifying dominant themes within the dataset.

The analysis used a minimum occurrence threshold of 5 for keyword inclusion, ensuring that only sufficiently frequent and relevant terms were visualized. The normalization of the co-occurrence network was performed using the association strength method, as implemented by default in

VOSviewer. Clustering was generated using the software's built-in modularity-based algorithm with a resolution parameter set to 1.00. No minimum cluster size restriction was imposed, allowing the natural structure of the network to emerge. To assess the robustness of the network structure, a sensitivity analysis was conducted by testing a lower occurrence threshold (minimum = 3) and applying fractional counting. The resulting network demonstrated comparable cluster structures and keyword relationships, indicating that the findings are stable across different parameter configurations.

To improve the consistency and accuracy of keyword co-occurrence analysis, a keyword normalization process was conducted. This involved merging synonymous terms, standardizing variations in spelling (e.g., "TPCK" to "TPACK"), and grouping related concepts under common labels (e.g., "e-learning" and "synchronous online classrooms" under "online learning"). The keyword cleaning and normalization dictionary is provided in the supplementary file.

2.4 Citation Analysis

Citation analysis was based on raw citation counts as indexed by Scopus on 15 November 2025. With regards to the identification of influential publications, only documents with 100 or more citations were included in the highly cited document analysis. Self-citations were not excluded, as the analysis aimed to capture overall scholarly influence within the field.

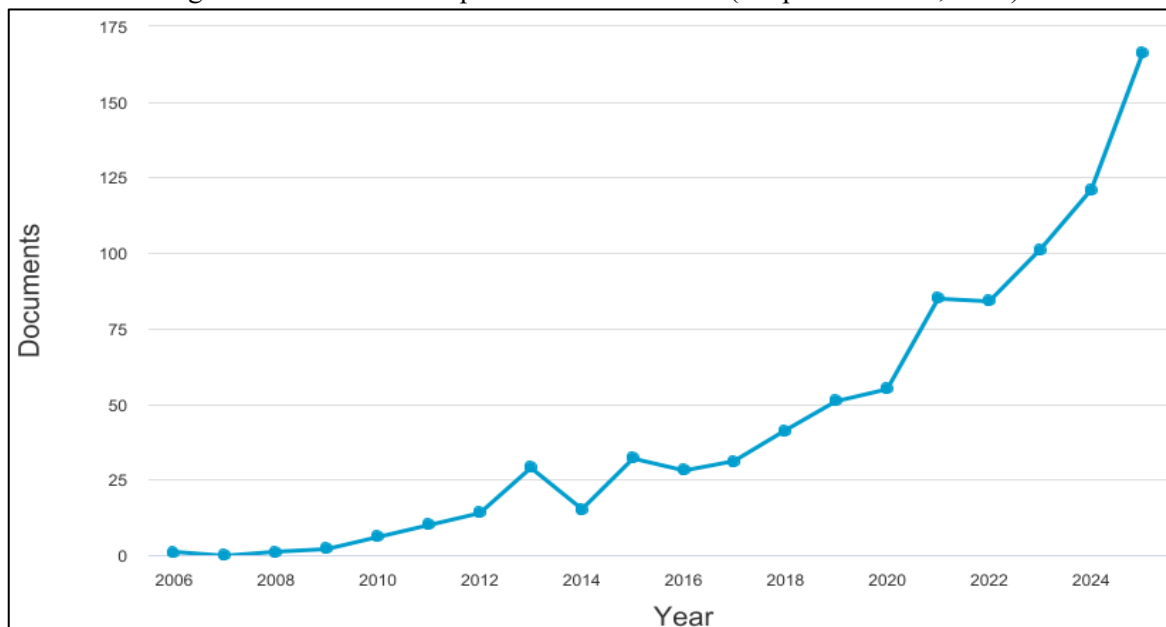
3. Findings

A total of 699 publications were reviewed by the researcher to be analyzed using Scopus descriptive analysis and VOS viewer for mapping the co-occurrence of keywords.

3.1 Publication Output by Year

The publication output retrieved from the Scopus database is shown below:

Figure 1. Publication Output from 2008 to 2025 (Scopus Database, 2025)



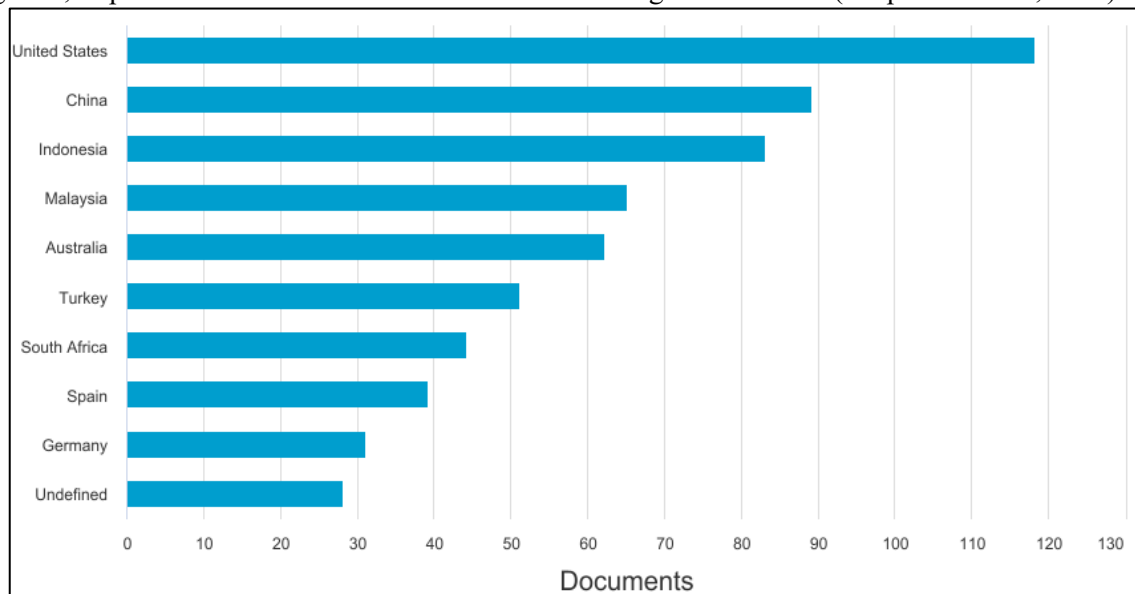
The annual publication trend shows a gradual increase (Figure 1) in TPACK-related studies in higher education from 2008 to 2025, configured from a total of 699 documents. In the early period (2008–2011), only few (less than 30) documents were published, followed by a small growth beginning around 2012. After a small peak in 2013 and a drop in 2014, the number of publications continued to grow

steadily. A more obvious increase appears from 2018 onward, with publication counts rising each year and reaching their highest levels between 2023 and 2025 with more than 100 documents each. Overall, the data indicates a consistent upward trend in publication output across the 18-year period.

3.2 Publication Documents by Country

The analysis shows that publication activity is led by a small group of highly productive countries (Figure 2). The United States records the highest number of documents, followed by China and Indonesia, indicating strong TPACK research engagement in these regions.

Figure 2, Top 10 Countries of TPACK Research Within Higher Education (Scopus database, 2025)

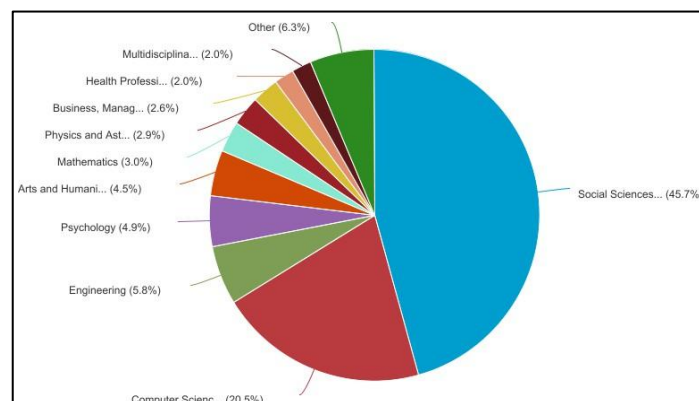


Other countries such as Australia, Malaysia, Turkey, and South Africa also contribute an extensive volume of publications. European countries including Germany and Spain appear in the average, while the Philippines records the lowest output among the top contributing nations. Overall, the distribution implies that research in this area is concentrated mainly in North America and Asia, with contributions from selected regions in Europe, Africa, and Southeast Asia.

3.3 Documents by Subject Area

The distribution of published documents based on subject area is shown in the Figure 3.

Figure 3. The Top 10 Subject Areas for Publication of TPACK Research within Higher Education Context Over the Last Two Decades (Scopus Database, 2025)



The distribution of subject-area (Figure 4) indicates that research on TPACK in higher education is predominantly anchored within the Social Sciences (45.7%), which account for almost half of all publications in the dataset. This reflects the strong positioning of TPACK within fields related to education, teaching, and learning research. Computer Science (21%) forms the second-largest group, highlighting the framework's close alignment with technology integration, digital systems, and instructional technology development. Other contributions emerge from Engineering (6.4%) and Psychology (5.3%), suggesting interest from disciplines concerned with technical, user behaviour, and cognitive aspects of technology-supported learning. Arts and Humanities (4%), as well as Business and Management (2.3%), also appear as meaningful contributors, indicating that TPACK has been applied in diverse instructional contexts beyond traditional teacher education. Smaller yet notable proportions of studies derive from Mathematics, Health Professions, Physics, and Multidisciplinary subject areas, demonstrating that the framework has gradually extended into domain-specific teaching practices. Overall, these patterns show that while the topic remains rooted primarily in social science and technology-related studies, it has attracted growing attention across a range of complementary academic fields.

3.4 Documents per Year by Source

Next, the subsequent bibliometric analysis is the top 5 source of TPACK research publications, shown in Figure 4:

Figure 4. The Top 5 Sources of TPACK Research Within Higher Education Context (Scopus database, 2025)

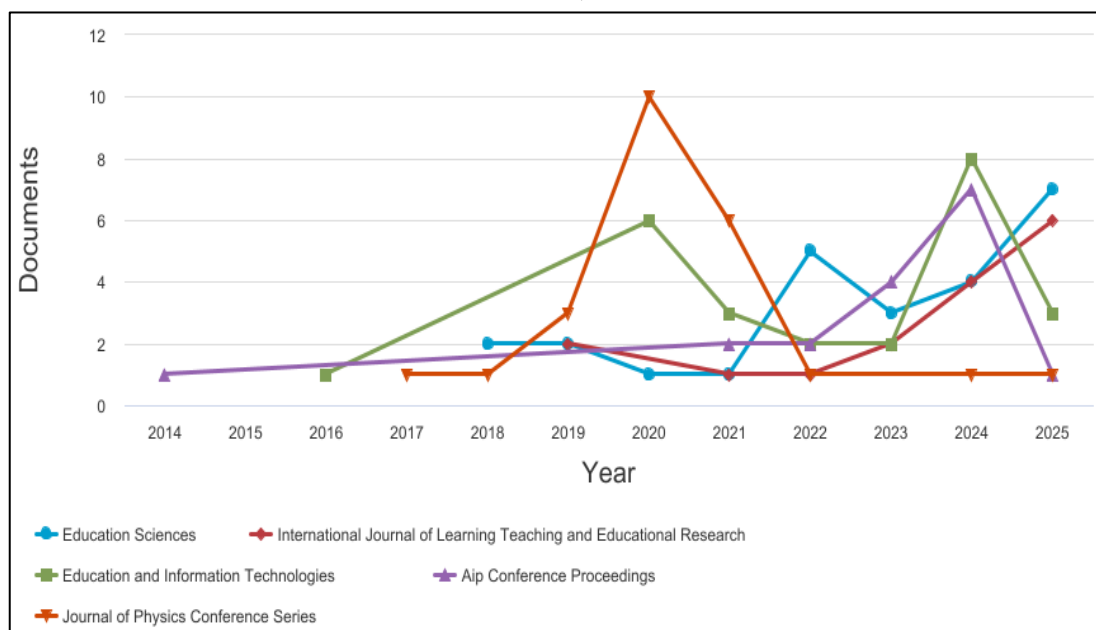


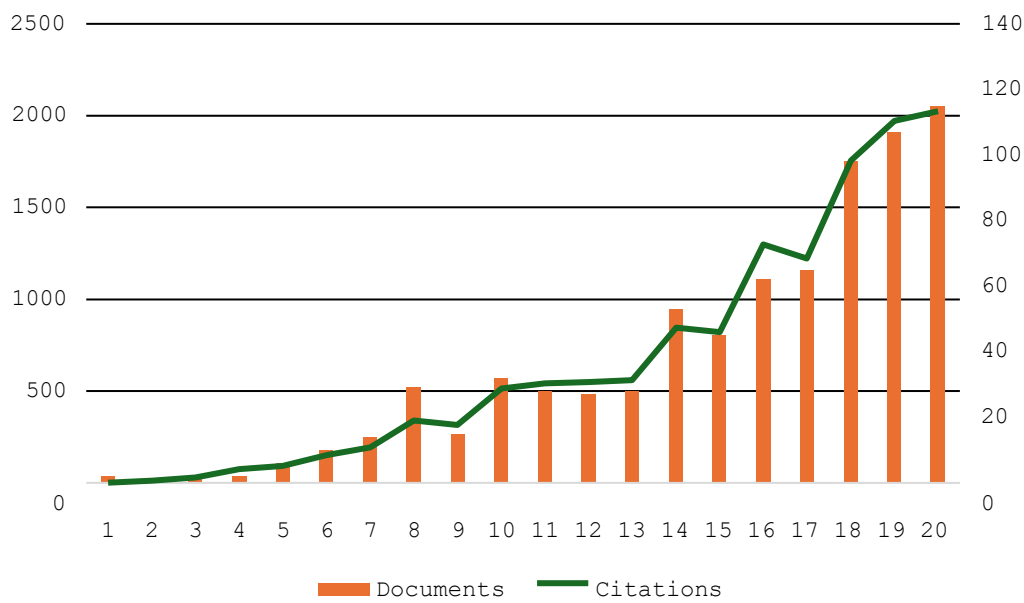
Figure 4 presents the distribution of publications across the most active journals over time. *Education and Information Technologies* showed the strongest and most consistent growth, beginning with a single publication in 2016 and increasing to a peak of eight documents in 2024. *Education Sciences* also demonstrated rising contributions, with noticeable peaks in 2022 and 2025. *Cogent Education* exhibited moderate but steady output, generally contributing between one and six documents annually. Conference-based sources such as the *AIP Conference Proceedings* displayed occasional publication patterns, with notable activity in 2021–2024 but reduced output in 2025. *Journal of Physics Conference Series* demonstrated spiked in 2020 with 10 documents but drop significantly two years after. Overall, the trend indicates that while a few core journals drive most of the TPACK publications, contributions

are subtle across both journals and conference proceedings, reflecting the framework's relevance to multiple academic communities within higher education.

3.5 Citations Output

The following bibliometric analysis discovers the citations output for TPACK research publications in higher education context:

Figure 5. Citation Trends for TPACK Research Publications (Scopus database, 2025)



The analysis of publication and citation trends recorded from 2011 to 2025 shows a clear upward trajectory in TPACK-related research within higher education. As illustrated in Figure 5, the number of documents increased gradually in the early years, with less than 10 publications annually between 2011 and 2015. A steady rise began in 2016, reaching 29 documents in 2018. The growth became more pronounced from 2020 (41 documents) onwards, aligning with global digital transformation in higher education. Publication output in 2025 pointed out with 108 documents, marking the highest annual contribution within the dataset. Citations followed a similar upward pattern, increasing slowly in the early years and then accelerating significantly from 2018 onward, reaching approximately 2,250 citations in 2025. To substantiate the observed growth trend, the compound annual growth rate (CAGR) was calculated for the period 2018–2025. The annual growth rate for the citation documents growth within five years between 2018 (39) to 2025 (112) is 23.47 percent, representing sustainable growth of TPACK research citations (Figure 5) during this period. This parallel increase suggests not only expansion in research output but also sustained scholarly influence and visibility of TPACK studies over time.

Next, we continue analyzing the most cited articles of TPACK research to see research interests, agendas and scopes that received attention from scholars toward TPACK. Only articles with more than 100 citations included in this list, while the remaining publications in the dataset were excluded as they received fewer than 100 citations. The list (16 documents) of highly cited documents (Table 3) was manually verified to ensure metadata accuracy. Records with similar titles were cross-checked using DOIs, author information, and source details to confirm that each entry represented a distinct publication. Minor discrepancies from Scopus interface rankings occur due to database update timing and duplicate documents were then verified. Thus, the rank orders for the most cited articles have been authenticated with the exported CSV format metadata.

Table 3. Documents with more than 100 citations

No.	Document title	Source	Author(s)	Citation
1.	Factors influencing preservice teachers' intention to use technology: TPACK, teacher self-efficacy, and Technology Acceptance Model	Educational Technology and Society	Joo, YJ; Park, S; Lim, E	413
2.	Profiling teachers' readiness for online teaching and learning in higher education: Who's ready?	Computers in Human Behavior	Scherer, R; Siddiq, F; Howard, SK; Tondeur, J.	350
3.	Teacher Educators' Use of Digital Tools and Needs for Digital Competence in Higher Education	Journal of Digital Learning in Teacher Education	Amhag, L; Hellström, L; Stigmar, M.	271
4.	The effects of online professional development on higher education teachers' beliefs and intentions towards learning facilitation and technology	Teaching and Teacher Education	Rienties, B; Brouwer, N; Lygo-Baker, S	218
5.	Preparing preservice teachers for self-regulated learning in the context of technological pedagogical content knowledge	Learning and Instruction	Kramarski, B; Michalsky, T	181
6.	Using digital technology to promote higher education learning: The importance of different learning activities and their relations to learning outcomes	Journal of Research on Technology in Education	Wekerle, C; Daumiller, M; Kollar, I.	144
7.	Using TPACK as a framework to understand teacher candidates' technology integration decisions	Journal of Computer Assisted Learning	Graham, CR; Borup, J; Smith, NB	143
8.	Investigating the impact of teacher education strategies on preservice teachers' TPACK	British Journal of Educational Technology	Baran, E.; Canbazoglu Bilici, S; Albayrak Sari, A; Tondeur, J	136
9.	Learning to teach online: Measuring the influence of faculty development training on teaching effectiveness through a TPACK lens	Internet and Higher Education	Brinkley-Etzkorn, KE	133
10.	Investigating pre-service teachers' acceptance of Web 2.0 technologies in their future teaching: a Chinese perspective	Interactive Learning Environments	Teo, T; Sang, G; Mei, B; Hoi, CKW	126
11.	Educating online student teachers to master professional digital competence: The TPACK-framework goes online	Computers and Education	Tømte, C; Enochsson, AB; Busqvist, U; Kårstein, A	125
12.	Mitigating the COVID-19 pandemic: a snapshot from Malaysia into the coping strategies for pre-service teachers' education	Journal of Education for Teaching	Mohamad Nasri, N; Husnin, H; Mahmud, SND; Halim, L	124
13.	Technology Integration in Higher Education During COVID-19: An Assessment of Online Teaching Competencies Through Technological Pedagogical Content Knowledge Model	Frontiers in Psychology	Akram, H; Yingxiu, Y; Al-Adwan, AS; Alkhalifah, A	111
14.	Preservice Teachers and Self-Assessing Digital Competence	Journal of Educational Computing Research	Maderick, J.A.; Zhang, S.; Hartley, K.; Marchand, G.	104

15.	The role of TPACK in affecting pre-service language teachers' ICT integration during teaching practices: Indonesian context	Education and Information Technologies	Habibi, A; Yusop, FD; Razak, RA	103
16.	Developing technological pedagogical content knowledge in pre-service mathematics teachers through collaborative design	Australasian Journal of Educational Technology	Agyei, DD; Voogt, J	101

*Citation counts as of November 15, 2025

The citation analysis identified most influential publications that received more than 100 citations within the TPACK and higher education research landscape. The most cited article was authored by Joo, Park, and Lim on pre-service teachers' intention to use technology (413 citations), followed by Scherer et al.'s study on online teaching readiness in higher education (350 citations). Other highly cited works reported on higher education teachers' digital competence needs (271 citations), the effects of online professional development (218 citations), and preservice teachers' self-regulated learning within a TPACK framework (181 citations). Additional publications addressed technology integration decisions, online teaching effectiveness, and faculty development (143–133 citations). The remaining article with citation counts between 101 and 126, focused on TPACK development strategies, pre-service teachers' acceptance of digital technologies, and the role of TPACK in subject-specific teaching such as language and mathematics. Overall, these highly cited studies largely concentrate on pre-service teacher preparation, digital competence, professional development, and technology integration in higher education contexts.

To complement total citation counts, citation impact was normalized using citations per year to account for differences in publication age. The researcher extracted the top 20 citations per year as presented in Table 4. The results indicate that several recent publications rank highly when adjusted for publication year. For example, studies by Wekerle, Akram and Yang demonstrate the higher annual citation impact compared to earlier publication documents. This suggests that while earlier studies dominate cumulative citation counts, more recent research has gained rapid scholarly attention, particularly in areas related to online learning, digital competence, and post-pandemic teaching practices. The inclusion of normalized citation metrics therefore provides a more balanced understanding of both established and emerging influential works in the field.

Table 4. Top 20 TPACK publications in higher education ranked by citations per year

Author (First)	Year	Citations	Citations/Year
Joo, Y.J.	2018	413	51.6
Scherer, R.	2019	350	50.0
Amhag, L.	2019	271	38.7
Wekerle, C.	2022	144	36.0
Akram, H.	2021	111	22.2
Mohamad Nasri, N.	2020	124	20.7
Baran, E.	2019	136	19.4
Habibi et al.	2020	103	17.2
Brinkley-Etz Korn, K.	2018	137	17.1
Rienties et al.	2013	218	16.8
Mpungose	2020	96	16.0
Teo, T.	2018	126	15.8
Yang et al.	2025	144	14
Tømte et al.	2015	125	11.4
Kramarski, B.	2010	181	11.3
Madriaga et al.	2016	104	10.4
Graham et al.	2012	143	10.2
Herring	2016	90	9.0

This indicates that TPACK research in higher education has evolved into a densely connected knowledge domain with strong conceptual integration. Rather than forming isolated thematic silos, the network demonstrates substantial overlap across clusters, suggesting that research topics such as digital competence, online learning, and instructional design are increasingly interconnected. These metrics indicate that TPACK research in higher education exhibits a high level of conceptual integration, with strong interconnections among key themes. Central nodes such as “TPACK,” “technology integration,” and “higher education,” play a significant role in connecting multiple thematic clusters within the network. While modularity and advanced centrality measures were not explicitly calculated, the clustering structure generated by VOSviewer and the observed link patterns provide sufficient evidence of the network’s thematic organization and intellectual cohesion.

The presence of 60 clusters further reflects the diversification of the field, where emerging topics and niche areas coexist alongside core themes. This pattern may indicate a transition from a relatively unified theoretical framework toward a more pluralistic and interdisciplinary research landscape. However, the continued centrality of core constructs such as TPACK suggests that the field remains anchored in its foundational framework, supporting a cumulative rather than fragmented development trajectory.

The visualization also shows a complex network connecting TPACK with higher education, forming a major thematic cluster (red). At the node level, the keyword “higher education” demonstrated 48 links with a total link strength of 147, indicating that it is moderately well-connected within the network. This suggests that while higher education serves as an important contextual anchor for literature, it is not the dominant conceptual driver of the network. Instead, it functions as a bridge term that links TPACK with various thematic areas such as teaching practices, digital competence, and technology integration. The strong links between higher education, students, and e-learning also highlight the framework’s application in studies exploring online, blended, and digitally supported learning environments.

A second major cluster (green) connects TPACK to curricula, engineering education, content knowledge, and teacher professional development, suggesting wider disciplinary applications outside teacher preparation programs. In the meantime, new themes such as artificial intelligence, generative AI, digital transformation, and learning design appear less connection but growing (blue and yellow clusters). This shows that TPACK study is starting to look more at new advance digital technologies.

Overall, the VOSviewer network structure proves that TPACK research in higher education is well-established and increasingly expanded. It also reveals the transforming environment for technology integration in higher education as the research areas on AI-enhanced learning, digital literacy, and curriculum innovation are gaining importance and emerging.

4. Discussion

4.1 Interpretation and comparison with prior reviews

The two-decade time frame (2006–2025) provides an important lens for interpreting the evolution of TPACK research in higher education. Since the framework was introduced by Mishra & Koehler (2006), the observed increase in publication output reflects its gradual transition from a conceptual model to a widely applied framework in digitally mediated teaching contexts. The sustained growth, particularly after 2018, can be linked to broader structural shifts in higher education, including increased institutional investment in digital learning infrastructure, expansion of online and blended learning programmes, and policy-driven digital transformation agendas (Kallou & Kikilia, 2021; Menon & Suresh, 2022).

The increase in publications after 2020 is likely associated with the COVID-19 pandemic (Chen & Hsu, 2022), which forced universities globally to rapidly adopt online and hybrid teaching models. This shift intensified the need to understand how educators integrate technology with pedagogy and content, positioning TPACK as a relevant analytical framework. While this study observes a strong growth trend, it should be interpreted cautiously as a correlation rather than a direct causal relationship, as other factors such as increased indexing coverage and publication accessibility may also contribute to this pattern.

In terms of geographical distribution, the dominance of the United States can be attributed to its early contribution to TPACK's conceptual development and its well-established research infrastructure in educational technology. The strong representation of China reflects national-level investments in digital education and large-scale higher-education expansion (Knox, 2020). Meanwhile, the prominence of Indonesia suggests increasing regional engagement with technology-enhanced teaching, possibly driven by national digitalisation initiatives and growing emphasis on higher-education reform (Aditya et al., 2020; Novianti & Nurlaelawati, 2019; Sofwan et al., 2021). These patterns indicate that TPACK research is influenced not only by theoretical interest but also by policy, funding priorities, and institutional capacity.

Compared with earlier reviews that focused primarily on school education (Shabira et al., 2024) or general contexts (Suprpto et al., 2021; Zou et al., 2022), this study highlights that TPACK research in higher education is both expanding and diversifying. However, the moderate connectivity of higher education as a keyword within the network suggests that the field remains conceptually anchored in general TPACK constructs, with limited theoretical differentiation specific to tertiary education contexts. This indicates a need for more context-sensitive research that explicitly addresses the unique pedagogical and institutional characteristics of higher education.

4.2 Practical Implications for Higher-Education Policy and Teacher Education

The findings of this review offer several practical implications for higher-education policy and academic development. First, universities should implement structured, institution-wide professional development programmes grounded in TPACK rather than relying on isolated or short-term training initiatives. Academic development units may prioritise blended TPACK training workshops tailored to disciplinary needs (e.g., STEM vs humanities), establish cross-faculty communities of practice, and incorporate pre- and post-assessment of TPACK using validated instruments to evaluate training effectiveness. Structured, sustained support rather than short-term workshops can better support educators in integrating technology meaningfully into pedagogy and content. Maipita et al. (2023) suggested that support from the institution is valuable for TPACK competency that will lead to high teachers' performance. Moreover, in today's digital world, university teachers are also required to have creative and innovative ways of working (Shafait & Huang, 2023) to improve their students' performance.

Second, teacher education and academic development programmes should embed TPACK-oriented competencies within curriculum design, assessment strategies, and teaching practice. This is particularly important in higher education, where many academics enter the profession without formal pedagogical training. Institutions may implement mentoring schemes, peer-observation frameworks, and collaborative course design initiatives to support educators in applying TPACK within authentic teaching contexts. Embedding TPACK within academic development programmes can therefore strengthen educators' confidence and competence in implementing technology-enhanced teaching strategies in sustainable ways.

Third, at the institutional level, universities should align digital teaching competence with performance evaluation and promotion systems. Recognising innovative teaching practices through promotion criteria, teaching awards, and workload allocation can incentivise sustained pedagogical innovation.

Additionally, institutional support mechanisms, including access to digital tools, instructional design support, and funding for teaching innovation should be systematically integrated into academic development strategies.

Collectively, these measures can support a more sustainable and contextually relevant integration of TPACK in higher education, moving beyond individual competence toward institutional transformation.

4.3 Future Research

Based on the identified findings, this review proposes several directions for future research. First, more longitudinal and explanatory studies are needed to examine how TPACK develops over time among higher-education educators. Future studies could address questions such as: *How does sustained institutional support influence the evolution of educators' TPACK across academic career stages?* Longitudinal survey designs or panel studies would be particularly suitable. In higher education, where formal pedagogical training is not usually necessary for academic positions, institutional environments can significantly influence how educators engage with technology-integrated teaching practices. Investigating TPACK development longitudinally across career stages helps determine when and how support mechanisms are most effective. Further determination based on different forms of support are needed for early-career, mid-career, and senior academics can be explored. This kind of research would assist universities design targeted academic development policies that promote continuous pedagogical innovation rather than one-off training interventions. Second, qualitative and design-based research remains underrepresented. In-depth case studies and design-based research could explore *how educators enact TPACK in authentic higher-education contexts*, particularly within digitally intensive disciplines such as tourism, engineering, and health sciences. The genuine beneficial output from various context in higher education might help the university stakeholders improve the curriculum, teaching methods, assessments and program outcomes.

4.4 Limitations

This study has several limitations that should be considered when interpreting the findings. First, the analysis is based solely on the Scopus database, which may introduce coverage bias, as relevant publications indexed in other databases such as Web of Science or Google Scholar were not included. Second, the restriction to English-language publications may exclude important contributions from non-English-speaking regions, potentially underrepresenting global research diversity.

Third, the inclusion of journal articles and conference papers may result in uneven representation across disciplines, as some fields rely more heavily on conference publications than others. Fourth, citation-based analyses are subject to citation-age bias, where older publications have had more time to accumulate citations compared to recent studies. Although citations per year were used to mitigate this issue, some bias may still remain.

Future research could address these limitations by incorporating multiple databases, including multilingual sources, and applying more advanced normalization techniques to further enhance the robustness of bibliometric findings.

5. Conclusion

This bibliometric review provides a comprehensive overview of two decades of TPACK scholarship in higher education, revealing a field that has grown steadily in figures, diversity, and global access since 2006. The analysis shows increasing research productivity, especially from 2018 onward, reflecting the rising importance of digital pedagogy, online learning, and digital applications in today's digital

education environment. Findings also demonstrate that TPACK-related studies are secured primarily within the social sciences but have expanded into STEM, vocational education, psychology, and other discipline domains. High citation works continue to shape the core understanding of teacher readiness, technology acceptance, and professional competence. On the other hand, more recent publications increasingly highlight pedagogical digital competence, innovative teaching behavior, and organizational support, projecting TPACK research to be more integrative with these domains. Overall, the review highlights how TPACK has evolved from a conceptual model to a widely applied framework guiding more empirical research, instructional innovation, and policy discussions across higher education.

6. Data and Code Availability

The dataset used in this study, including the DOI list, keyword normalization dictionary, and VOSviewer input files, is provided as supplementary material.

7. Funding

This research received no external funding.

8. Conflict of Interest

The authors declare no conflict of interest.

References

- Aditya, B. R., Ferdiana, R., & Kusumawardani, S. S. (2020). The Study of the Barriers to Digital Transformation in Higher Education: A Preliminary Investigation in Indonesia. *Proceedings - 2020 6th International Conference on Science and Technology, ICST 2020*. <https://doi.org/10.1109/ICST50505.2020.9732809>
- Ali, W. (2020). Online and Remote Learning in Higher Education Institutes: A Necessity in light of COVID-19 Pandemic. *Higher Education Studies*, 10(3), 16. <https://doi.org/10.5539/hes.v10n3p16>
- Barrios-del-Ángel, A., ... L. J.-F.-E., & 2026, undefined. (2026). Digital Literacy Research for Teaching Excellence in Higher Education: A Bibliometric Perspective. *Mdpi.Com*, 14(1). <https://doi.org/10.1002/REV3.70145>
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L. A., & Otto, A. (2022). Teachers' digital competencies in higher education: a systematic literature review. In *International Journal of Educational Technology in Higher Education* (Vol. 19, Number 1). <https://doi.org/10.1186/s41239-021-00312-8>
- Cabero, J., & Barroso, J. (2016). ICT teacher training: a view of the TPACK model / *Formación del profesorado en TIC: una visión del modelo TPACK*. *Cultura y Educación*, 28(3), 633–663. <https://doi.org/10.1080/11356405.2016.1203526>
- Chen, Y.-J., & Hsu, L. (2022). A comparative research on teachers' knowledge in five Asia-Pacific countries in the COVID-19 pandemic: The case of tourism and hospitality education. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.911182>
- Doğru, M., Çelik, M., & Satar, C. (2020). Bibliometric profile of scientific studies on pedagogical content knowledge. *Socialinis Ugdymas*, 51(1), 6–23. <https://doi.org/10.15823/su.2019.51.1>
- Hsu, Y.-S., Yeh, Y.-F., & Wu, H.-K. (2015). The TPACK-P Framework for Science Teachers in a Practical Teaching Context. In *Development of Science Teachers' TPACK* (pp. 17–32). Springer Singapore. https://doi.org/10.1007/978-981-287-441-2_2
- Kallou, S., & Kikilia, A. (2021). A transformative educational framework in tourism higher education through digital technologies during the COVID-19 pandemic. *Advances in Mobile Learning Educational Research*, 1(2), 37–47. <https://doi.org/10.25082/AMLER.2021.01.005>
- Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, 45(3), 298–311. <https://doi.org/10.1080/17439884.2020.1754236>
- Lindfors, M., Pettersson, F., & Olofsson, A. D. (2021). Conditions for professional digital competence: the teacher educators' view. *Education Inquiry*, 12(4). <https://doi.org/10.1080/20004508.2021.1890936>
- Maipita, I., Dongoran, F. R., Syah, D. H., & Sagala, G. H. (2023). TPACK, Organizational Support, and Technostress in Explaining Teacher Performance During Fully Online Learning. *Journal of Information Technology Education: Research*, 22. <https://doi.org/10.28945/5069>
- Mayasari, V., Indyastuti, D. L., & Daryono, D. (2024). Pedagogic competence of lecturers with non-educational backgrounds in the challenges of 21st century learning. *Salud, Ciencia y Tecnología - Serie de Conferencias*, 3. <https://doi.org/10.56294/SCTCONF2024.996>

- Menon, S., & Suresh, M. (2022). Enablers of technology agility in higher education. *International Journal of Information and Learning Technology*, 39(2). <https://doi.org/10.1108/IJILT-07-2021-0107>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. In *Teachers College Record* (Vol. 108, Number 6). <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Novianti, N., & Nurlaelawati, I. (2019). Pedagogical Competence Development of University Teachers with Non-Education Background: The Case of a Large University of Education in Indonesia. *International Journal of Education*, 11(2), 172. <https://doi.org/10.17509/ije.v11i2.15711>
- Omar, A. M., & Abdullahi, M. O. (2024). A bibliometric analysis of sustainable digital transformation in developing countries' higher education. *Frontiers in Education*, 9. <https://doi.org/10.3389/FEDUC.2024.1441644/FULL>
- Rad, D., Learning, G. R.-S. E. and S., & 2025, undefined. (2025). A Bibliometric Analysis of Smart Learning Environments Under the Digital Pedagogy Paradigm. *Igi-Global.Com*, 30(2), 2547–2575. <https://doi.org/10.1007/S10639-024-12899-2>
- Redecker, C., & Punie, Y. (2017). European framework for the digital competence of educators: DigCompEdu. In *Joint Research Centre (JRC) Science for Policy report*. <https://doi.org/10.2760/159770>
- Saidi, Z. A. & Abd Aziz, M. F. (2025). Empowering student leadership in higher education: A literature review on key competencies for effective leadership. *Journal of Institutional Research South East Asia*, 23(1), 376-399.
- Sanchez, M. A. (2020). University E-readiness for the Digital Transformation: the Case of Universidad Nacional del Sur. *Revista Gestão & Tecnologia*, 20(2). <https://doi.org/10.20397/2177-6652/2020.v20i2.1848>
- Santoveña-Casal, S., & Lopez, S. R. (2024). Mapping of digital pedagogies in higher education. *Education and Information Technologies*, (29), 2437–2458. <https://doi.org/10.1007/s10639-023-11888-1>
- Shabira, Q., Baharudin, & Yanti, Y. (2024). Mapping the Literature of Technological Pedagogical and Content Knowledge (TPACK) in Elementary Education: A Bibliometric Review. *Jurnal Penelitian Pendidikan IPA*, 10(9), 631–643. <https://doi.org/10.29303/jppipa.v10i9.8731>
- Shafait, Z., & Huang, J. (2023). Exploring the Nexus of Emotional Intelligence and University Performance: An Investigation Through Perceived Organizational Support and Innovative Work Behavior. *Psychology Research and Behavior Management*, 16. <https://doi.org/10.2147/PRBM.S422194>
- Sofwan, M., Pratama, R., Muhaimin, M., Yusnaidar, Y., Mukminin, A., & Habibi, A. (2021). Contribution of technology innovation acceptance and organizational innovation climate on innovative teaching behavior with ICT in Indonesian education. *Qwerty*, 16(1). <https://doi.org/10.30557/QW000035>
- Suprpto, N., Sukarmin, S., Puspitawati, R. P., Erman, E., Savitri, D., Ku, C.-H., & Mubarak, H. (2021). Research trend on TPACK through bibliometric analysis (2015-2019). *ERIC*, 10(4), 1375–1385. <https://doi.org/10.11591/ijere.v10i4.22062>

Zou, D., Huang, X., Kohnke, L., Chen, X., Cheng, G., & Xie, H. (2022). Research trend on Technological Pedagogical Content Knowledge (TPACK) through bibliometric analysis (2015-2021). *Apjbhe.Segi.Edu.My*, 27(8), 10585–10609. <https://doi.org/10.1007/s10639-022-10991-z>

Beyond Badges: A Bibliometric Exploration of Micro-Credential Research in Higher Education (2015 – 2025)

ABSTRACT

Authors: Nur Hazimah Nor Hashim & Su Luan Wong*
Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
 *Corresponding author: suluan@upm.edu.my

Received: 31/12/2025 | **First Revision:** 14/1/2026
Second Revision: 15/4/2026: | **Accepted:** 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Micro-credentials have emerged as a key innovation in higher education, offering short, flexible, and skills-oriented certifications that support employability and lifelong learning. Despite their growing adoption, persistent challenges related to quality assurance, interoperability, and governance continue to fragment practice across pedagogical, technological, and policy domains. This study maps the intellectual landscape of micro-credential research between 2015 and 2025 through a bibliometric analysis of 571 Scopus-indexed publications, following rigorous screening and eligibility procedures. Citation indicators and network visualisations were generated using Bibliomagika® (Ahmi, 2024) and VOSviewer. The findings reveal sustained and accelerating growth in scholarly output, particularly after 2019, indicating increasing institutional engagement with flexible credentialing models. However, research productivity and influence remain unevenly distributed, with the United States dominating output and citation impact, while countries such as Australia demonstrate higher citation efficiency. Thematic analysis identifies five major clusters: pedagogical foundations and competency-based learning, learner engagement and assessment practices, digital credentialing and technological infrastructure, instructional innovation and educational ecosystems, and alternative credential pathways and employability. Importantly, the convergence of keyword and co-authorship analyses shows that the field is not only conceptually segmented but also structurally fragmented, with collaboration concentrated within clusters and limited integration across domains. The findings indicate that micro-credential research is pedagogically anchored, technologically emerging, institutionally evolving, and economically relevant yet weakly integrated with employability discourse. This configuration reflects a field developing through parallel trajectories rather than as a cohesive system. By integrating Diffusion of Innovation (DOI) and Resource-Based View (RBV) perspectives, this study advances understanding of micro-credentials as a socio-technical-policy framework shaped by asynchronous diffusion processes and unequal resource configurations. While the field is transitioning from early exploration towards structured consolidation, its continued advancement depends on strengthening integration across pedagogical, technological, and policy domains. The study provides evidence-informed insights for educators, institutions, and policymakers to support the development of coherent, scalable, and inclusive micro-credential systems aligned with Sustainable Development Goal 4 (United Nations, 2023).

Keywords: Bibliometric analysis, employability, higher education, lifelong learning, micro-credentials, research fragmentation, socio-technical-policy framework, thematic clustering

1. Introduction

Over the past decade, higher education systems have undergone significant transformation driven by globalisation, technological advancements, and increasing demand for flexible learning pathways. Ahsan et al. (2023) and Ngoc Ha et al. (2025) noted that these changes have transformed learner expectations and institutional structures, positioning micro-credentials as a prominent educational innovation. Also referred to as modular and competency-based credentials, micro-credentials are designed to enhance employability, promote lifelong learning, and adapt to evolving labour market demands (Ahsan et al., 2023; Ngoc Ha et al., 2025). Micro-credentials are typically offered online and are used to verify focused skills and competencies that meet industry demands and lifelong learning goals (Oliver, 2019; Tamoliūnė et al., 2023).

The growing significance of micro-credentials was further reinforced by Varadarajan et al. (2023), who pointed out that the COVID-19 pandemic made their importance more apparent. The rapid adoption of online learning during this period exposed the limitations of traditional, time-bound qualifications, particularly their inflexibility in addressing rapidly evolving skills needs. As a result, micro-credentials are being positioned by both governments and universities as strategic tools to broaden access to education and support large-scale reskilling efforts. Reflecting on this shift, national policy frameworks and institutional innovation agendas have begun to include these credentials (Moodie & Wheelahan, 2021; Ngoc Ha et al., 2025). Despite this growing momentum, concerns persist regarding their credibility (Oliver, 2019), scalability, and long-term sustainability (Ahsan et al., 2023; Ngoc Ha et al., 2025).

From a research perspective, the body of literature on micro-credentials has expanded rapidly; however, it remains conceptually and methodologically fragmented. Existing studies vary widely in scope, dataset size, and analytical approach, resulting in limited cumulative knowledge development (Ahsan et al., 2023; Ngoc et al., 2022; Tamoliūnė et al., 2023). While some studies focus on specific themes, technologies, or institutional contexts, others adopt descriptive or narrative approaches that do not systematically capture the structural evolution of the field (Cudney et al., 2022; Selvaratnam & Sankey, 2021). Consequently, there is limited understanding of how micro-credential research has developed over time, how knowledge production is distributed across regions and institutions, and which themes and publications have exerted the greatest intellectual influence (Narong, 2025; Rullyana et al., 2025). This fragmentation is further reflected in the dispersion of micro-credential research across multiple domains, including pedagogy, technology, and policy, with limited integrative synthesis linking these dimensions (Martín-Ramos et al., 2025; Vordenberg et al., 2024). Ahsan et al. (2023) and Tamoliūnė et al. (2023) further highlighted that such fragmentation constrains the development of coherent theoretical explanations and weakens the capacity of research to inform evidence-based policy and institutional strategies.

Although bibliometric approaches have been widely applied in related domains such as Massive Open Online Courses (MOOCs) and Open Educational Resources (OER), where they have been used to map intellectual structures, thematic evolution, and collaboration networks (Irwanto et al., 2023; Krishnan et al., 2025), their application to micro-credential research remains limited and methodologically inconsistent. Existing bibliometric studies often rely on single databases, small or narrowly defined datasets, or descriptive trend analyses, without integrating temporal patterns, thematic structures, collaboration networks, and intellectual influence within a unified analytical framework (Narong, 2025; Yüksel, 2025; Strielkowski, 2025; Johan et al., 2025; Rullyana et al., 2025). As a result, the intellectual structure and global development of micro-credential research remain insufficiently understood.

In order to address these limitations, this study adopts a bibliometric approach to systematically examine the evolution of micro-credential research between 2015 and 2025. Using the Scopus database for its extensive coverage of peer-reviewed literature, the study integrates performance analysis and network

visualisation techniques through Bibliomagika® (Ahmi, 2024) and VOSviewer (Van Eck & Waltman, 2014) to analyse publication trends, research productivity, thematic clusters, and intellectual influence. This approach enables a comprehensive and replicable mapping of the field's development.

To strengthen explanatory power, this study is guided by Diffusion of Innovation (DOI) theory and the Resource-Based View (RBV) (Guo & Huang, 2024). DOI is employed to interpret temporal patterns of research development, particularly the emergence, acceleration, and consolidation of scholarly output (Guidolin & Manfredi, 2022; Takahashi et al., 2024). RBV, in turn, is used to explain disparities in research productivity and influence across countries and institutions, based on differences in resources such as funding, technological infrastructure, and research capacity (Vargo et al., 2020). These theoretical lenses are integrated within a socio-technical-policy perspective, which conceptualises micro-credential research as an ecosystem shaped by the interaction of pedagogical practices, technological infrastructures, and policy frameworks.

Accordingly, this study addresses the following research questions:

1. What are the publication trends and evolutionary patterns of micro-credential research between 2015 and 2025?
2. What are the patterns and disparities in research output and impact across countries, institutions, and authors in micro-credential research?
3. What are the prevailing thematic clusters in micro-credential research?
4. Which publications have exerted the greatest intellectual influence on the field?

By addressing these questions, this study provides an empirically grounded and theory-informed understanding of the intellectual landscape of micro-credential research. It contributes to the field by offering an integrated analytical perspective that moves beyond descriptive accounts towards a systematic explanation of its development, thereby supporting more informed research, policy, and institutional practices.

2. Literature Review

2.1 Conceptualising Micro-Credentials as a Multi-Dimensional Construct

Micro-credentials are commonly defined as short, targeted certifications that validate specific knowledge or skill sets outside traditional degree programmes (Abdul Halim et al., 2024; UNESCO, 2022). Their modular and stackable nature enables learners to accumulate smaller units of learning that can contribute towards broader qualifications, thereby supporting flexible and lifelong learning pathways (Ngoc Ha et al., 2025; Tamoliūnė et al., 2023). In addition, micro-credentials are often delivered through digital platforms and may take various forms, including alternative certificates and digital badges (Ahsan et al., 2023; Oliver, 2019).

Despite this general agreement, the literature reveals considerable conceptual variation. Some studies categorise micro-credentials under broader constructs such as “alternative credentials” (Abdul Halim et al., 2024), while others associate them closely with digital badges and platform-based certification systems (Santally et al., 2024). This variation extends beyond terminology and reflects deeper epistemological differences across disciplines. Educational research tends to conceptualise micro-credentials as pedagogical tools grounded in competency-based and learner-centred approaches, whereas technology-oriented studies frame them as digital artefacts embedded within platform ecosystems and verification infrastructures such as blockchain. Policy-oriented scholarship, in contrast, positions micro-credentials within governance, quality assurance, and labour market alignment frameworks (UNESCO, 2022).

Regional perspectives further illustrate this diversity. Asian scholarship often emphasises workforce reskilling and employability, European research focuses on policy integration and standardisation, while

Anglo-American studies highlight micro-credentials as mechanisms for enhancing labour market competitiveness (Ngoc Ha et al., 2025; Santally et al., 2024). While this diversity demonstrates the adaptability of micro-credentials across contexts, it also contributes to conceptual fragmentation, complicating cross-study comparison and limiting coherent policy alignment (UNESCO, 2022). These variations reinforce the need for integrative analytical approaches that capture the interaction between pedagogical, technological, and policy dimensions.

2.2 Micro-Credentials in Higher Education: Opportunities, Tensions, and Structural Inequalities

The integration of micro-credentials into higher education has generated both optimism and critical debate. On one hand, micro-credentials are recognised for their potential to enhance employability, widen participation, and provide flexible pathways that complement traditional qualifications (Ahsan, 2023; Santally et al., 2024). Evidence from sectors such as technology, healthcare, and business demonstrates that they can support professional development and increase learner motivation (Varadarajan et al., 2023). However, the literature consistently highlights persistent challenges related to quality assurance, standardisation, and recognition. Concerns regarding employer acceptance, comparability across institutions, and long-term credibility remain central to ongoing debates (Moodie & Wheelahan, 2021; Ngoc Ha et al., 2025). These issues are not merely operational but reflect deeper structural tensions within higher education systems. Adoption patterns further demonstrate uneven development across regions. In Australia and the European Union, structured frameworks have been established to guide implementation, whereas in many other contexts micro-credentials remain at pilot or exploratory stages with inconsistent certification processes (Lang, 2023; Šarcevic, 2024). This uneven development reflects broader structural inequalities in global higher education.

From a Resource-Based View (RBV) perspective, these disparities can be attributed to differences in access to critical resources such as funding, technological infrastructure, institutional capacity, and policy support (Vargo et al., 2020). Institutions with stronger resource endowments are better positioned to implement and scale micro-credentials, thereby reinforcing their dominance in both practice and research output. Complementing this, Diffusion of Innovation (DOI) theory explains these variations as different stages of adoption, where early adopters shape global discourse while others remain in emerging or transitional phases (Guidolin & Manfredi, 2022; Takahashi et al., 2024). These perspectives collectively highlight that the development of micro-credentials is structurally conditioned by both resource availability and diffusion dynamics.

2.3 Fragmentation and Limitations in Bibliometric Research

Despite the rapid expansion of micro-credential research, the field lacks a systematically mapped intellectual structure. Much of the existing literature remains descriptive, focusing on opportunities and challenges without examining patterns of knowledge production, collaboration networks, or intellectual influence (Ahsan, 2023; Tamoliūnė et al., 2023). Recent bibliometric and review-based studies have attempted to examine the development of micro-credential research; however, a comprehensive and integrative understanding remains limited due to variations in scope, dataset size, and analytical approaches. Existing studies range from small-scale analyses focusing on specific domains, such as artificial intelligence in Open and Distance Learning (Harizan & Ally, 2025), to larger-scale investigations identifying thematic clusters and publication trends (Strielkowski, 2025; Narong, 2025). While these studies provide useful insights, they are often constrained by reliance on single databases, geographically concentrated datasets, or predominantly descriptive analytical approaches (Rullyana et al., 2025; Yüksel, 2025; Johan et al., 2025). Even large-scale analyses, such as Savita et al. (2025), remain limited by their dependence on author-provided keywords and limited contextual interpretation.

Notably, existing literature is dominated by systematic and narrative reviews rather than comprehensive bibliometric analyses, indicating an imbalance between qualitative synthesis and quantitative mapping approaches (Alharthi, 2025; Yildirim, 2025; Warnars, 2024; Ahsan, 2023; Alsobhi, 2023; Tamoliūnė et al., 2023). These limitations highlight the absence of a methodologically consistent and integrative

bibliometric study capable of capturing temporal evolution, thematic structures, collaboration networks, and intellectual influence within a unified analytical framework. As shown in Table 1, these studies vary considerably in dataset size, database selection, and analytical techniques, further reinforcing the fragmented nature of the field. This methodological inconsistency constrains cumulative knowledge development and limits the ability to generate theory-driven insights. Without systematic mapping, it remains difficult to identify dominant research domains, emerging trends, and structural disparities, thereby reducing the practical relevance of existing research for policy and institutional decision-making.

Table 1. Summary of Selected Bibliometric Studies on Micro-Credentials

Study	Scope & Dataset	Methodology	Key Findings	Limitations	Contribution to Current Study
Harizan & Ally (2025)	46 articles examining AI integration in micro-credentials within Open and Distance Learning (ODL)	Hybrid bibliometric and thematic analysis	AI integration remains fragmented, limiting field consolidation; identifies key research gaps and future directions	Narrow focus on AI and ODL; small dataset (n = 46)	Bridges theory–practice gap; informs industry applications and future research directions
Strielkowski (2025)	608 Scopus-indexed publications (1992–2025)	Bibliometric network analysis using VOSviewer	Identifies six thematic clusters, including motivation, alternative credentialing, lifelong learning, AI-enhanced learning, teacher professional development, and equity	Single database (Scopus); geographical concentration in developed regions	Highlights role of micro-credentials in AI-driven higher education; emphasises need for digital infrastructure
Narong (2025)	401 Scopus-indexed documents	Systematic analysis with descriptive statistics, keyword co-occurrence, and science mapping	Field remains in early stage; strong concentration in developed countries; focus on STEM, healthcare, and blockchain	Scopus-only dataset; limited maturity of knowledge base	Establishes post-pandemic research baseline and thematic benchmarks
Rullyana et al. (2025)	551 publications (Scopus and	PRISMA, performance analysis, and	Rapid growth after 2018; dominant	Emphasis on trend mapping;	Identifies global trends and underexplored

	WoS, 2014–2024)	science mapping (VOSviewer, RStudio)	contributions from the United States; themes include blockchain, gamification, and personalised learning	limited conceptual integration	areas such as information literacy and curriculum integration
Yüksel (2025)	85 WoS-indexed articles	Bibliometric and keyword co-occurrence analysis using VOSviewer	Strong association with higher education, employability, and digital badges; identifies five thematic clusters	Limited dataset; restricted to English WoS publications	Provides higher education–focused insights and identifies key research areas
Johan et al. (2025)	85 Scopus-indexed journal articles (2015–2023)	Bibliometric and performance analysis using VOSviewer	Social sciences dominate the field (54.2%); major contributors include the US, Australia, and Canada; limited collaboration networks	Scopus-only dataset; narrow disciplinary focus; inclusion of in-press articles	Highlights lack of research collaboration and need for interdisciplinary engagement
Savita et al. (2025)	3,743 WoS-indexed publications (1970–2024)	Bibliometric and text analytics (Python/pyBibX), word cloud, and co-occurrence analysis	Significant growth between 2014–2020; key themes include motivation, digital badges, and ethical data concerns	Reliance on author keywords; limited contextual and regional analysis	Provides longitudinal mapping and highlights need for ethical governance and impact evaluation

Note(s): MCs = micro-credentials; WoS = Web of Science; ODL = Open and Distance Learning.

Accordingly, bibliometric analysis offers a robust and replicable approach to address these limitations by enabling the systematic identification of intellectual structures, thematic patterns, and collaboration networks (Van Eck & Waltman, 2014).

2.4 Conceptual Framework: A Socio-Technical–Policy Perspective

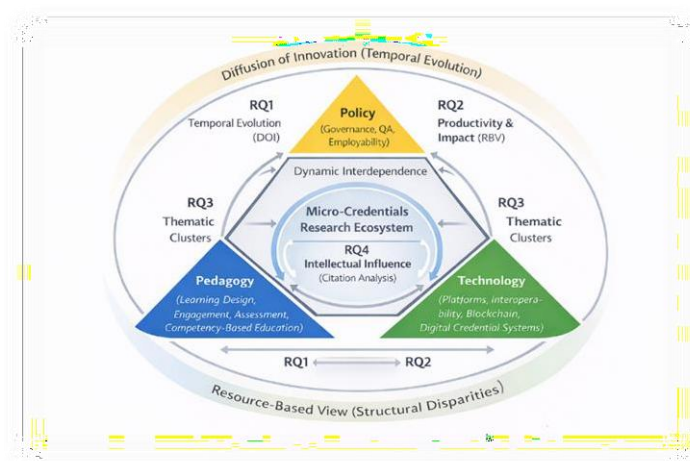
In order to address the fragmentation identified in the literature, this study adopts a socio-technical–policy framework that conceptualises micro-credential research as an interconnected ecosystem shaped

by the interaction of pedagogical, technological, and policy dimensions. This perspective recognises that micro-credentials operate as a system of interdependent components rather than isolated elements. The pedagogical dimension encompasses learning design, learner engagement, and assessment practices that underpin micro-credential development. The technological dimension focuses on digital platforms, interoperability, and credential verification systems, including emerging technologies such as blockchain. The policy dimension includes governance structures, quality assurance mechanisms, and employability alignment that influence institutional adoption and recognition. These dimensions are mutually constitutive and often develop unevenly. Technological advancements may outpace policy frameworks, while pedagogical innovation may be constrained by institutional governance structures. This interdependence highlights the need for an integrative analytical approach that captures both alignment and imbalance within the system.

This framework is integrated with Diffusion of Innovation (DOI) and the Resource-Based View (RBV) to enhance its explanatory capacity. DOI accounts for the temporal evolution of micro-credential research, particularly the progression from emergence to consolidation phases (Guidolin & Manfredi, 2022; Takahashi et al., 2024), while RBV explains disparities in research productivity and influence by linking them to variations in resource availability and institutional capacity (Vargo et al., 2020). These perspectives provide a comprehensive analytical lens. While the socio-technical-policy framework identifies the key domains shaping the field, DOI explains how the field evolves over time, and RBV explains why certain actors dominate this evolution. This integrated approach enables the study to move beyond descriptive mapping towards theory-driven explanation, directly supporting the research objectives.

This integrated conceptualisation is illustrated in Figure 1, which presents micro-credential research as a dynamic socio-technical ecosystem shaped by the interdependence of pedagogical, technological, and policy dimensions. The figure visually demonstrates how these domains interact to form the intellectual structure of the field, while also highlighting their uneven development. In particular, the outer layers represent the theoretical lenses guiding this study, where Diffusion of Innovation (DOI) explains temporal evolution (RQ1), and the Resource-Based View (RBV) accounts for disparities in research productivity and impact (RQ2). The framework further aligns thematic clustering (RQ3) and intellectual influence (RQ4) with the internal structure of the ecosystem. As such, Figure 1 serves as an analytical scaffold guiding the interpretation of bibliometric findings.

Figure 1. Socio-Technical-Policy Analytical Framework for Micro-Credential Research.



Note (s): The framework conceptualises micro-credential research as a dynamic ecosystem shaped by the interaction of pedagogical, technological, and policy dimensions. These domains are analysed through bibliometric techniques aligned with four research questions, examining temporal evolution,

productivity disparities, thematic structures, and intellectual influence. Diffusion of Innovation (DOI) informs the interpretation of temporal growth patterns, while Resource-Based View (RBV) explains disparities in institutional and geographical research productivity. The framework functions as an analytical scaffold guiding data interpretation rather than a descriptive model.

3. Methodology

3.1 Database Selection

This study employed the Scopus database as the primary data source for bibliometric analysis (Ahsan et al., 2023; Irwanto et al., 2023; Krishnan et al., 2025). Scopus was selected for its comprehensive coverage of peer-reviewed journals, conference proceedings, books, and book chapters across education, computer science, and interdisciplinary domains (Irwanto et al., 2023; Krishnan et al., 2025). Compared with Web of Science or Google Scholar, Scopus provides broader indexing and consistent metadata formats, which are essential for reliable keyword extraction and citation mapping (Ahsan et al., 2023; Irwanto et al., 2023). Moreover, the choice of Scopus aligns with prior bibliometric studies on digital learning and credentialing (Ahsan et al., 2023; Irwanto et al., 2023). The data were extracted on 21 August 2025 at 10:28pm to ensure transparency and to acknowledge the dynamic nature of the database.

3.2 Search Strategy

After determining the database, a comprehensive search strategy was developed using Boolean operators and truncation to capture core concepts and related terminology associated with micro-credentials. The final search string applied in Scopus was as follows:

```
TITLE-ABS-KEY ( "micro-credential*" OR "micro credential*" OR "digital badge*"
OR "alternative credential*" OR "open badge*" OR "digital certificate*" )
AND PUBYEAR > 2014 AND PUBYEAR < 2026
AND ( LIMIT-TO ( LANGUAGE , "English" ) )
AND ( EXCLUDE ( PREFNAMEAUID , "undefined" ) )
```

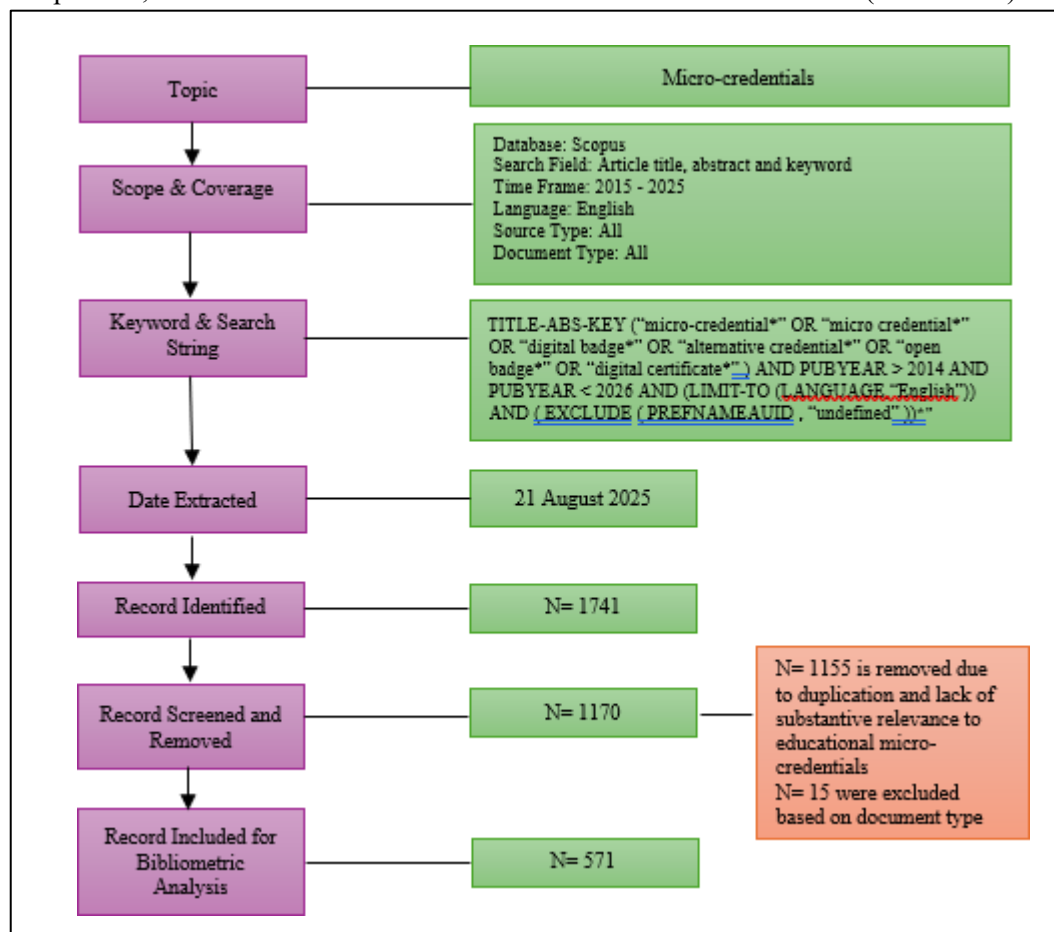
The TITLE-ABS-KEY field was selected to balance inclusiveness and precision by retrieving records in which the key concepts appeared in titles, abstracts, or author keywords. In addition, the filter EXCLUDE (PREFNAMEAUID, "undefined") was applied to remove records with missing or undefined author identifiers. This decision was made to improve the accuracy of author-level and co-authorship analyses. A sensitivity check further indicated that this exclusion did not materially alter the composition of the dataset or the overall thematic structure. At the preliminary stage, terms such as nanodegree, micro-course, and virtual badge were also tested. However, these terms were then excluded because they frequently retrieved records associated with commercial training platforms, standalone short courses, or gamification rewards rather than formal educational credentialing frameworks. Their inclusion substantially increased the retrieval of vendor-driven or non-academic records and reduced conceptual specificity. Additional sensitivity checks were conducted by temporarily including these terms; however, they introduced a substantial proportion of non-academic records and did not alter the core thematic structure, thereby supporting their exclusion. The process of identifying and screening the publications is presented in Figure 2, which follows the methodological approach of Zakaria et al. (2021) and Zhang et al. (2025), and is structured in accordance with the PRISMA guidelines (Moher et al., 2009).

3.3 Screening and Eligibility

Following the search process, the initial Scopus search retrieved 1,741 records. The records were exported in CSV format with full bibliographic fields, including authors, affiliations, abstracts,

keywords, references, and citation counts. The screening process was then conducted manually in Microsoft Excel, while Bibliomagika® was used to support dataset validation and preprocessing. Next, the records were screened according to predefined inclusion and exclusion criteria. During this stage, 1,155 records were removed following duplicate identification and lacked substantive relevance to educational micro-credentials. In addition, 15 records were excluded based on document type, comprising 8 notes, 4 editorials, 1 errata, 1 letter, and 1 short survey. After the screening and eligibility assessment, the final dataset consisted of 571 publications, including journal articles (276), conference papers (164), book chapters (110), reviews (14), and books (7). Figure 2 illustrates the final set of 571 publications included in the bibliometric dataset. To ensure conceptual consistency, studies were included only if they explicitly examined micro-credentials, digital badges, open badges, or alternative credentials in educational contexts; addressed pedagogical, institutional, learner, or assessment dimensions; were peer-reviewed outputs indexed in Scopus; and provided either empirical or conceptual contributions. Conversely, studies were excluded if they focused exclusively on blockchain infrastructure, cryptographic certificates, or cybersecurity systems without pedagogical relevance; examined credentialing systems in non-educational contexts; treated badges merely as gamification tools; presented purely technical system designs without educational application; were duplicate records; or referred to credentials only superficially without substantive engagement with alternative credentialing in education.

Figure 2. Flow diagram of search strategy showing the publication identification and screening process, the inclusion of 571 records in the final bibliometric dataset (2015–2025).



Source: Moher et al. (2009), Zakaria et al. (2021), Zhang et al. (2025)

3.4 Time Frame and Scope

Once the eligible records had been identified, the scope of the analysis was defined. The study covered publications from 2015 to 2025, thereby capturing a decade of emergence and consolidation in micro-credential research. Consistent with the study's eligibility criteria, only English-language, peer-reviewed outputs were retained, while non-academic records such as trade publications, editorials, errata, and short surveys were excluded. The final dataset of 571 publications exceeds the minimum commonly recommended for bibliometric analysis and is therefore sufficient to support stable network visualisation and citation-based interpretation (Rogers et al., 2020).

3.5 Data Cleaning and Harmonisation

After the final dataset had been established, metadata cleaning and harmonisation were conducted to improve consistency and reliability. This process was carried out using Bibliomagika® (Ahmi, 2024) and OpenRefine (Ahmi, 2023). Institutional names were standardised to the parent organisation level. For example, faculty- or campus-level variants such as "Purdue Univ." were harmonised to "Purdue University". Affiliations were consolidated at the institutional rather than faculty or city level, and country names were standardised to recognised sovereign entities. For instance, "Park University" was carefully distinguished from "University Park," as the former refers to a private university in Missouri, United States, whereas the latter denotes a geographical location associated with Pennsylvania State University.

OpenRefine was used to cluster and merge variant author keywords to ensure conceptual consistency. Examples of harmonisation included "competencies" to "competency," "microcredential" to "micro-credentials," "HE" to "higher education," and "organization" to "organisation." The harmonisation process followed a consistent set of rules, including singular-plural standardisation, British-American spelling alignment, abbreviation expansion, and the merging of semantically equivalent terms. In addition, Bibliomagika Split was applied to disaggregate compound metadata fields.

3.6 Data Analysis Tools and Procedures

After cleaning and harmonising the metadata, the bibliometric analysis was conducted using Bibliomagika® and VOSviewer. Bibliomagika® (Ahmi, 2024) was used to generate descriptive statistics and citation indicators, including publication counts, document-type distributions, subject-area frequencies, total citations, h-index, g-index, and citations per publication. In addition, citation normalisation was incorporated by reporting both total citations (TC) and citations per year (C/Y) to account for differences in publication age. Due to limitations in Scopus export regarding self-citation filtering, citation counts were retained as reported; however, this limitation is acknowledged in the discussion section.

VOSviewer (version 1.6.20) was used to construct and visualise bibliometric networks, specifically keyword co-occurrence and co-authorship networks. For the keyword co-occurrence analysis, the following parameters were applied: full counting method, association strength normalisation, and a minimum keyword occurrence threshold of 5. A total of 1,108 keywords were identified, of which 57 met the inclusion threshold. The minimum cluster size was set to 1, and clustering was performed using the default VOSviewer algorithm. These settings ensured that only frequently occurring and conceptually meaningful keywords were retained, thereby enhancing the robustness and interpretability of the network structure. Based on these criteria, five distinct thematic clusters were generated. For the co-authorship analysis, the unit of analysis was set at the author level. The analysis employed the full counting method with a minimum threshold of one document per author. A total of 1,429 authors were identified, all of whom met the inclusion threshold. The minimum cluster size was set to 1, and clustering was conducted using the default VOSviewer settings. Based on these parameters, six clusters

were generated, representing distinct collaboration groups within the micro-credential research landscape.

To assess the robustness of network clustering, standard VOSviewer parameters (association strength normalisation and default clustering resolution) were applied. While formal modularity and silhouette coefficients are not directly generated in VOSviewer, cluster coherence was assessed through interpretive validation by examining the semantic consistency of grouped keywords. While temporal network segmentation (e.g., dividing into sub-periods) can provide additional insights, this study captures temporal evolution through longitudinal publication trends (RQ1) and keyword emergence patterns. Future research may extend this analysis using time-sliced network visualisation.

Configuration settings and screenshots are provided in the supplementary materials. The combined use of Bibliomagika® and VOSviewer enabled triangulation across descriptive, citation-based, and network analyses, thereby enhancing the robustness of the findings (Ahsan et al., 2023; Van Eck & Waltman, 2014). All steps were documented carefully to ensure transparency and support replication.

4. Findings and Discussion

The analysis was guided by four research questions:

Table 2. Mapping of Research Questions to Data Sources, Analytical Methods, and Bibliometric Outputs

Research Question	Data Source	Analytical Method	Tool(s) Used	Key Parameters & Outputs (Tables/Figures)
RQ1: Publication trends and evolutionary patterns (2015–2025)	Scopus metadata (publication year, citation counts)	Longitudinal bibliometric analysis of publication output and citation dynamics to identify growth phases and evolutionary patterns	Bibliomagika®	Time span: 2015–2025; indicators: TP, TC, C/P, h-index. Outputs: Table 3 (main information); Figure 3 (annual publication and citation trends; foundation, acceleration, consolidation phases)
RQ2: Patterns and disparities in research output and impact across countries, institutions, and authors	Author affiliations, country data, and citation metrics	Productivity and impact analysis at author, institutional, and country levels; comparative analysis of citation performance to identify disparities	Bibliomagika®	Full counting method; minimum threshold: ≥ 5 publications (authors), ≥ 10 (institutions), ≥ 9 (countries). Outputs: Table 4 (authors), Table 5 (institutions), Table 6 (countries)
RQ3: Prevailing thematic clusters in micro-credential research	Author keywords; subject area classification	Keyword co-occurrence analysis and clustering to identify dominant research themes; disciplinary distribution analysis to contextualise thematic structure	VOSviewer; Bibliomagika®	Full counting; minimum keyword occurrence = 5; clustering resolution (default). Outputs: Table 7 (subject area distribution); Figure 4 (keyword co-occurrence network and thematic clusters); Table 8 (keyword clusters)
RQ4: Publications with the greatest intellectual influence	Citation data (Scopus); author relationships	Citation analysis to identify highly influential publications; co-authorship network analysis to examine	Bibliomagika®; VOSviewer	Ranking criteria: total citations (TC), citations per year (C/Y), recency. Network visualisation parameters (association strength). Outputs: Table 9

		intellectual collaboration patterns		(Top 20 highly cited articles); Figure 5 (co-authorship network)

This alignment ensured coherence between dataset construction, analytical procedures, and interpretation of results, providing a rigorous and replicable bibliometric roadmap of micro-credential research.

RQ1: What are the publication trends and evolutionary patterns of micro-credential research between 2015 and 2025?

The first research question aimed to examine how micro-credential research has developed over time. This study combines bibliometric indicators (Table 3) with publication and citation trends shown in Figure 3. Together, these provide a clear picture of how the field has grown, how much attention it has received, and how it is evolving.

Table 3. Main Information

Main Information	Data
Publication Years	2015 - 2025
Total Publications	571
Citable Year	12
Number of Contributing Authors	1730
Number of Cited Papers	462
Total Citations	6,879
Citation per Paper	12.05
Citation per Cited Paper	14.89
Citation per Year	687.90
Citation per Author	3.98
Author per Paper	3.03
Citation sum within h-Core	5,466
h-index	42
g-index	62
m-index	3.500

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

The bibliometric profile presented in Table 3 shows that micro-credential research is growing steadily but is still developing as a field. The dataset includes 571 publications produced between 2015 and 2025, involving 1,730 contributing authors. This reflects a growing and increasingly collaborative research community. The citation indicators suggest that the field is gaining attention but has not yet fully matured. The average citation per paper (12.05) and per cited paper (14.89) indicate a moderate level of influence, which is common for newer research areas. At the same time, the relatively high citation per year (687.90) shows that interest in this topic has increased, especially in recent years. The overall impact of the field can also be seen through the h-index (42), g-index (62), and m-index (3.500). These values suggest that while many studies are being published, a smaller number of key publications are receiving more attention and shaping the direction of the field. In other words, the field is growing, but influence is still concentrated in selected works. Authorship patterns also support this observation. On average, there are 3.03 authors per paper, indicating that research in this area is often collaborative. This is expected because micro-credentials involve different areas such as education, technology, and policy, which often require input from multiple disciplines.

To further address RQ1, Figure 3 illustrates how the field has changed over time. The overall trend shows a gradual increase in publications, rising from approximately 27 publications in 2015 to a peak of around 129 publications in 2024, before a slight decline to about 61 publications in 2025. This pattern indicates steady growth, with more noticeable acceleration after 2019 indicating that micro-credentials have become more widely discussed, particularly following the shift towards online learning during the COVID-19 period.

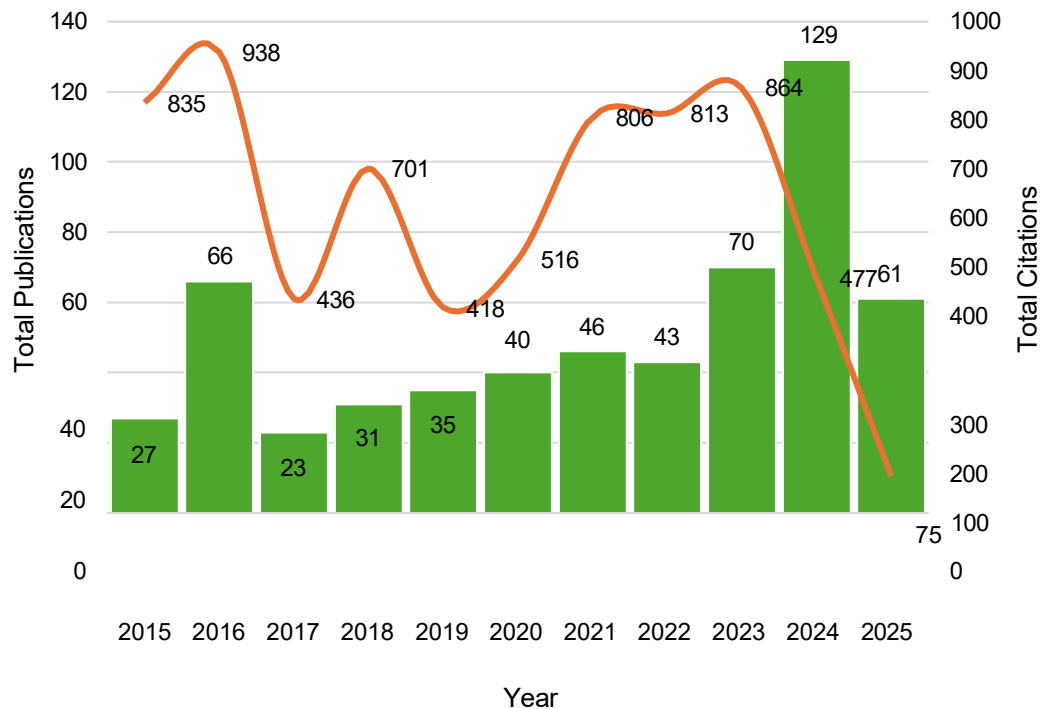


Figure 3. Total Publications and Citations by Year

The phases are categorised as foundation (2015–2018), acceleration (2019–2021), and consolidation (2022–2025).

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

Based on this trend, the development of the field can be understood in three phases, which broadly reflect the Diffusion of Innovation (DOI) pattern of gradual emergence, expansion, and consolidation. The first phase (foundation from 2015–2018) represents the early stage, where publication output remained relatively low, fluctuating between approximately 27 and 45 publications per year. Despite the lower volume, citation levels during this period were relatively high, indicating that early studies played an important role in shaping key concepts and directions in the field.

The second phase (acceleration from 2019–2021) shows a clearer and more consistent increase in publications, growing from about 35 publications in 2019 to 46 in 2021. During this period, research began to expand into different areas, and more scholars became interested in the topic. As a result, influence became more distributed across studies rather than being concentrated in a few early works.

The third phase (consolidation from 2022–2025) reflects a more rapid increase in publication output, rising from around 43 publications in 2022 to 70 in 2023, and reaching a peak in 2024 (approximately 129 publications). This suggests that micro-credentials have become a widely recognised topic across different disciplines. However, citation counts for recent years appear lower, which is expected because newer publications have had less time to accumulate citations.

Overall, the findings show a clear pattern of development. In the early stage, the field was shaped by a small number of influential studies. This was followed by a period of expansion, where more researchers contributed, and the topic gained wider attention. More recently, the field has grown rapidly, but

research has become more diverse and less centralised. This suggests that micro-credential research is

moving from an early stage of development towards a more established field. However, it is still in the process of strengthening its theoretical and conceptual foundations. While publication output continues to increase, the field has not yet reached full maturity, and there is still room for further development and integration.

RQ2: What are the patterns and disparities in research output and impact across countries, institutions, and authors in micro-credential research?

To address the second research question, examining the patterns and disparities in research output and impact across authors, institutions, and countries in micro-credential research, the analysis begins with the contributions of the most productive authors. The focus is placed on authors with a minimum of five publications, with key bibliometric indicators including total publications, cited papers, total citations, citation impact, and h, g, and m-indices examined in detail.

Table 4. Most Productive Authors with a minimum of five publications

Author's Name	Current Affiliation	Country	TP	NCP	TC	C/P	C/C P	h	g	m
Hickey, Daniel Thomas	Indiana University	United States	12	11	130	10.83	11.82	5	11	0.417
West, Richard Edward	Brigham Young University	United States	10	9	216	21.60	24.00	8	10	0.667
Davis, Katie	University of Washington	United States	7	7	200	28.57	28.57	6	7	0.500
Abramovich, Samuel	University at Buffalo	United States	7	5	74	10.57	14.80	3	7	0.273
Newby, Timothy James	Purdue University	United States	6	6	181	30.17	30.17	6	6	0.667
Jovanović, Jelena M.	University of Belgrade	Serbia	6	6	154	25.67	25.67	3	6	0.250
Devedžić, Vladan B.	University of Belgrade	Serbia	6	6	161	26.83	26.83	4	6	0.333
Cheng, Zui	Shenzhen University	China	6	6	160	26.67	26.67	6	6	0.667
Fanfarelli, Joseph R.	Marist College	United States	6	4	48	8.00	12.00	3	6	0.250
Horstman, Theresa	University of Washington	United States	5	4	26	5.20	6.50	2	5	0.286
McDaniel, Rudy	University of Central Florida	United States	5	5	59	11.80	11.80	4	5	0.333
McGreal, Rory	Athabasca University	Canada	5	4	200	40.00	50.00	4	5	0.800
Gamrat, Chris W.	Pennsylvania State University	United States	5	4	27	5.40	6.75	4	5	0.333
Tierney, Gavin	California State University	United States	5	4	26	5.20	6.50	2	5	0.286
Willis, James E.	Indiana University	United States	5	5	33	6.60	6.60	3	5	0.250

Note: TP=total number of publications; NCA=Number of contributing authors; NCP=number of cited publications; TC=total citations; C/P=average citations per publication; C/CP=average citations per cited publication; h=h-index; g=g-index; m=m-index.

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

At the author level, the results show that productivity and influence are not always aligned. For example, Hickey is the most productive author with 12 publications, but the average citation per paper (C/P =

10.83) is relatively moderate. In contrast, authors such as Davis ($C/P = 28.57$), Newby ($C/P = 30.17$), and McGreal ($C/P = 40.00$) have fewer publications but higher citation impact. This indicates that some authors contribute fewer but more influential studies. Hence, the findings suggest that influence in this field is concentrated among a smaller group of authors rather than evenly distributed.

A similar pattern can be observed at the institutional level (Table 5). Leading institutions, particularly from the United States, such as Indiana University, Purdue University, and the University of Washington, produce a high number of publications. However, high productivity does not always mean high impact. For instance, Indiana University shows strong output but lower citation impact ($C/P = 10.17$), while Curtin University has fewer publications but much higher impact ($C/P = 41.90$). Purdue University demonstrates a more balanced profile, combining both productivity and influence. This suggests that institutions play different roles, with some acting as major contributors in terms of volume, while others contribute fewer but more impactful studies.

Table 5. Most Productive Institutions with a Minimum of Ten Publications

Affiliation	Country	TP	NCA	NCP	TC	C/P	C/CP	<i>h</i>	<i>g</i>	<i>m</i>
Indiana University	United States	18	32	17	183	10.17	10.76	8	13	0.667
Purdue University	United States	16	40	14	301	18.81	21.50	10	16	0.833
University of Washington	United States	15	30	13	251	16.73	19.31	9	15	0.750
Pennsylvania State University	United States	14	34	11	196	14.00	17.82	6	14	0.500
Brigham Young University	United States	11	19	9	216	19.64	24.00	8	11	0.667
Curtin University	Australia	10	15	9	419	41.90	46.56	5	10	0.417

Note: TP=total number of publications; NCA=number of contributing authors; NCP=number of cited publications; TC=total citations; C/P=average citations per publication; C/CP=average citations per cited publication; *h*=*h*-index; *g*=*g*-index; *m*=*m*-index.

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

At the country level (Table 6), the results show clear differences in global participation and influence. The United States leads in both publication output (TP = 216) and total citations (TC = 2,857), indicating a dominant position in the field. However, when citation impact is considered, countries such as Australia ($C/P = 24.53$), Canada ($C/P = 22.04$), and New Zealand ($C/P = 19.64$) show higher average influence per publication. On the other hand, countries such as Malaysia ($C/P = 7.47$), China ($C/P = 5.53$), and Indonesia ($C/P = 2.89$) are increasing their publication output but still have lower citation impact. This indicates that while participation is becoming more global, influence remains uneven.

Table 6. Top countries with a minimum of 9 publications

Country	TP	NCA	NCP	TC	C/P	C/CP	<i>h</i>	<i>g</i>	<i>m</i>
United States	216	571	183	2857	13.23	15.61	26	53	2.167
Australia	55	122	51	1349	24.53	26.45	20	36	1.667
United Kingdom	40	86	34	390	9.75	11.47	11	19	0.917
Malaysia	30	128	22	224	7.47	10.18	6	14	0.545
Canada	28	50	25	617	22.04	24.68	9	24	0.750
Spain	22	57	17	256	11.64	15.06	7	16	0.778
Germany	20	48	16	289	14.45	18.06	7	17	0.636
China	17	53	14	94	5.53	6.71	6	9	0.500

Country	TP	NCA	NCP	TC	C/P	C/CP	<i>h</i>	<i>g</i>	<i>m</i>
Turkey	16	29	13	83	5.19	6.38	5	9	0.455
New Zealand	14	39	14	275	19.64	19.64	7	14	0.778
Ireland	14	43	11	236	16.86	21.45	7	14	1.000
Finland	13	38	10	163	12.54	16.30	6	12	0.667
Italy	11	30	8	43	3.91	5.38	4	6	0.500
South Africa	11	18	8	199	18.09	24.88	4	11	0.333
Indonesia	9	38	6	26	2.89	4.33	2	5	0.333
Austria	9	22	8	53	5.89	6.63	3	7	0.600

Note: TP=total number of publications; NCA=number of contributing authors; NCP=number of cited publications; TC=total citations; C/P=average citations per publication; C/CP=average citations per cited publication; *h*=*h*-index; and *g*=*g*-index.

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

To sum up, these findings point to a clear pattern in the structure of the field. A small group of countries and institutions contributes a large share of the research output, while influence is concentrated among selected authors and institutions. This creates a situation where some parts of the world play a more central role in shaping research directions, while others are still developing their presence. From a Resource-Based View (RBV) perspective, these differences can be explained by variations in access to resources such as funding, research infrastructure, and international collaboration networks. Countries and institutions with stronger resources are more likely to produce more research and gain higher visibility. In contrast, those with fewer resources may contribute to the field but face challenges in achieving similar levels of impact.

Overall, the findings show that micro-credential research is expanding globally, but it remains uneven in terms of productivity and influence. While more countries and institutions are participating, the field is still shaped by a relatively small number of dominant contributors. This suggests that although the field is growing, there is still a need for more balanced development and wider distribution of research influence.

RQ3. What are the prevailing themes in micro-credential research?

The prevailing themes in micro-credential research were examined using three approaches: subject-area distribution (Table 7), keyword co-occurrence analysis (Figure 4), and thematic cluster interpretation (Table 8). These methods provide an overview of both the disciplinary background and the main research themes in the field.

The subject-area analysis shows that most studies are located within the Social Sciences (73.73%), followed by Computer Science (39.75%) and Engineering (15.24%). This indicates that micro-credential research is mainly driven by educational and policy-related concerns, while technological disciplines play a supporting role. The presence of other areas such as Business, Psychology, and Arts and Humanities further demonstrates that the field is gradually expanding to include organisational and human-centred perspectives. This disciplinary distribution provides the foundation for the thematic structure observed in the keyword co-occurrence network (Figure 4), where pedagogical and educational concepts emerge as central nodes linking multiple research areas.

Table 7. Subject Area Distribution

No.	Subject Area	TP	(%)
1	Social Sciences	421	73.73%
2	Computer Science	227	39.75%

3	Engineering	87	15.24%
4	Business, Management and Accounting	54	9.46%
5	Psychology	24	4.20%
6	Arts and Humanities	22	3.85%
7	Decision Sciences	20	3.50%
8	Mathematics	19	3.33%
9	Medicine	18	3.15%
10	Nursing	13	2.28%
11	Economics, Econometrics and Finance	12	2.10%
12	Health Professions	10	1.75%
13	Energy	7	1.23%
14	Materials Science	7	1.23%
15	Physics and Astronomy	7	1.23%
16	Chemistry	5	0.88%
17	Environmental Science	4	0.70%
18	Multidisciplinary	3	0.53%
19	Pharmacology, Toxicology and Pharmaceutics	3	0.53%
20	Agricultural and Biological Sciences	2	0.35%
21	Biochemistry, Genetics and Molecular Biology	1	0.18%
22	Chemical Engineering	1	0.18%
23	Earth and Planetary Sciences	1	0.18%
24	Neuroscience	1	0.18%

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

The keyword co-occurrence network (Figure 4) reveals five main clusters that represent the key themes in the field. Central keywords such as micro-credentials and digital badges appear as highly connected nodes, showing that they link multiple areas of research. While the clusters are interconnected, they are not yet fully integrated, indicating that the field is still developing across different domains. These network patterns are further clarified in Table 8, which groups co-occurring keywords into distinct thematic clusters and explains their roles within the broader research landscape.

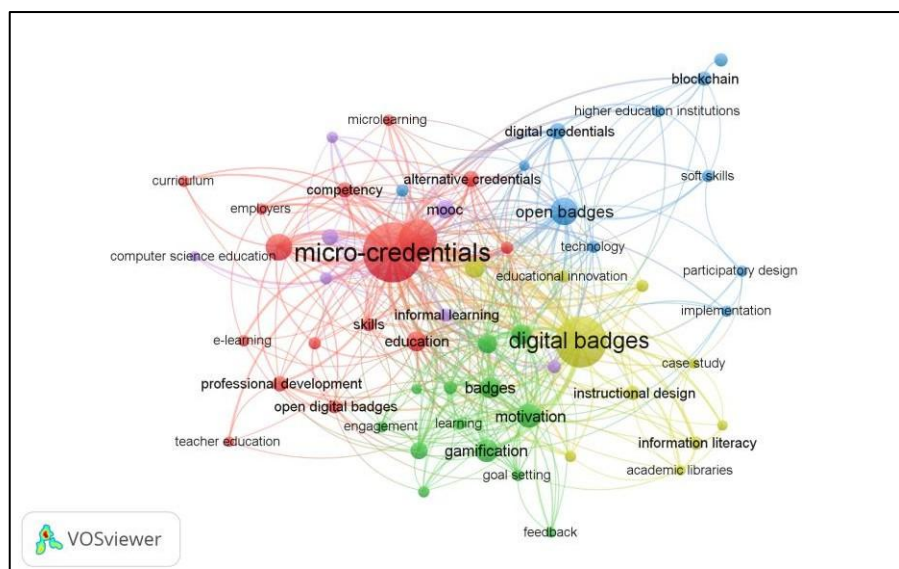


Figure 4. Network visualization of the co-occurrence analysis of the author keywords
Source: Generated by the author(s) using VOSviewer (van Eck & Waltman, 2014)

Table 8 provides a clearer summary of these five clusters by grouping related keywords into meaningful themes and explaining their roles within the field. The clusters show how micro-credential research is organised across different areas, including pedagogy, technology, and employability.

Table 8. Keyword Clusters in Micro-Credential Research (2015–2025)

Cluster	Colour	Theme Label	Representative Keywords	Thematic Interpretation
Cluster 1	Red	Pedagogical Foundations and Competency-Based Learning	Micro-credentials, higher education, lifelong learning, competency, skills, curriculum, professional development, e-learning, teacher education	Represents the core educational foundation of micro-credential research, emphasising competency-based learning, skills development, and curriculum integration within higher education. This cluster highlights how micro-credentials are conceptualised as flexible learning units supporting lifelong learning and workforce readiness.
Cluster 2	Green	Learner Engagement and Assessment Practices	Motivation, gamification, assessment, engagement, student engagement, feedback, learning, online learning	Focuses on learner-centred pedagogical processes, particularly engagement, motivation, and assessment strategies. This cluster reflects how micro-credentials are designed to enhance participation, self-regulation, and learning outcomes through interactive and performance-based approaches.
Cluster 3	Blue	Digital Credentialing and Technological Infrastructure	Open badges, digital credentials, blockchain, technology, implementation, higher education institutions, soft skills	Captures the technological dimension of micro-credentials, including digital credentialing systems, verification mechanisms, and platform-based implementation. The presence of blockchain indicates increasing attention to secure, portable, and interoperable credential systems.
Cluster 4	Yellow	Instructional Innovation and Educational Ecosystems	Digital badges, instructional design, information literacy, educational innovation, academic libraries, training, case study	Reflects applied educational practices and institutional innovation, particularly in instructional design and knowledge ecosystems such as academic libraries. This cluster highlights how micro-credentials are integrated into broader teaching and learning environments
Cluster 5	Purple	Alternative Credential Pathways and Employability	MOOC, employability, informal learning,	Represents flexible and alternative learning pathways linked to labour market outcomes. This cluster emphasises recognition,

			recognition, short courses, badging	employability, and non-traditional education routes, positioning micro-credentials within broader lifelong learning and workforce development frameworks.

From a socio-technical–policy perspective, the identified clusters can be interpreted as representing three interrelated dimensions of micro-credential research: pedagogical (learning and assessment), technological (digital infrastructure and credentialing systems), and institutional–policy (implementation, governance, and employability).

4.1 Cluster 1: Pedagogical Foundations and Competency-Based Learning

Situated within the pedagogical dimension of the socio-technical–policy framework and consistent with the dominance of Social Sciences in Table 7, this cluster forms the core educational foundation, with micro-credentials (occurrence = 199; total link strength = 307) acting as the dominant node, strongly connected to higher education, lifelong learning, and competency-based education. The prominence of these terms highlights that micro-credentials are primarily conceptualised as mechanisms for skills development, curriculum innovation, and continuous learning pathways within formal education systems. The inclusion of professional development, learning analytics, and microlearning further indicates a shift towards data-informed and modular pedagogical approaches, where learning is increasingly personalised and performance-driven. At the same time, the presence of employers and skills reflects a growing alignment between academic learning outcomes and labour market expectations, reinforcing the role of micro-credentials as bridges between education and employability. However, despite this strong pedagogical grounding, the comparatively lower occurrence of terms such as teacher education and curriculum suggests that instructional implementation and educator readiness remain underexplored, indicating a potential gap between conceptual frameworks and practical application.

4.2 Cluster 2: Learner Engagement and Assessment Practices

Also situated within the pedagogical dimension, this cluster focuses on learner engagement and assessment practices, with key terms such as *motivation* (occurrence = 28), *gamification* (26), *assessment* (20), and *engagement* (6). The co-occurrence of these terms highlights the increasing emphasis on enhancing learner participation and improving learning outcomes through interactive and technology-supported strategies. The integration of badges, credentials, and digital badging indicates that assessment is not only evaluative but also representational, where achievements are translated into visible and transferable credentials. This reflects a shift from traditional assessment models towards formative, continuous, and outcome-based evaluation systems. Moreover, the presence of *goal setting*, *feedback*, and *student engagement* indicates that learning in micro-credential environments is often framed within self-regulated learning paradigms, where learners actively monitor and direct their own progress. Nevertheless, the less strongly connected link strength between this cluster and technological infrastructure clusters implies that engagement strategies are not yet fully integrated with advanced digital systems, suggesting a partial disconnect between pedagogical innovation and technological capability.

4.3 Cluster 3: Digital Credentialing and Technological Infrastructure

Situated within the technological dimension and aligned with the presence of Computer Science and Engineering in Table 7, this cluster captures the digital infrastructure and credentialing systems, with *open badges* (occurrence = 38) and *digital credentials* as central nodes, supported by terms such as *blockchain*, *technology*, and *implementation*. The inclusion of *blockchain* reflects increasing interest in secure, verifiable, and decentralised credentialing systems, addressing issues of trust, authenticity, and portability. The presence of *higher education institutions*, *workforce development*, and *soft skills* demonstrates that digital credentialing systems are not only technical solutions but also institutional and economic tools, enabling the recognition of diverse competencies across educational and professional contexts. However, the relatively lower total link strength compared to Cluster 1 indicates that this technological domain remains less integrated with core pedagogical discourse, reinforcing the notion that technological advancement is progressing somewhat independently of instructional design and learning theory. Additionally, the inclusion of *participatory design* and *implementation* highlights the complexity of adopting these technologies, suggesting that institutions must navigate organisational, technical, and user-experience challenges when operationalising digital credential systems.

4.4 Cluster 4: Instructional Innovation and Educational Ecosystems

Positioned at the intersection of pedagogical and institutional dimensions, this cluster reflects the application of micro-credentials within educational ecosystems, with *digital badges* (occurrence = 143; total link strength = 257) and *instructional design* as central elements. The strong presence of *educational innovation*, *educational technology*, and *training* indicates that micro-credentials are increasingly embedded within structured teaching and learning frameworks. The inclusion of *information literacy*, *academic libraries*, and *library instruction* indicates that micro-credentials are being integrated into broader educational ecosystems, extending beyond traditional classrooms to support lifelong and interdisciplinary learning. Essentially, this cluster acts as a mediating layer between pedagogical theory (Cluster 1) and technological systems (Cluster 3), translating conceptual ideas into practical instructional strategies and institutional practices. However, the appearance of *case study* as a recurring term indicates that much of the research in this area remains context-specific and exploratory, with limited large-scale or generalisable evidence. This suggests that while implementation is growing, the field is still in a developmental and experimental phase.

4.5 Cluster 5: Alternative Credential Pathways and Employability

Reflecting the policy and labour market dimension, this cluster highlights the role of micro-credentials in employability and flexible learning pathways, with key terms such as *MOOC*, *employability*, *informal learning*, and *recognition*. These terms reflect the role of micro-credentials in supporting non-traditional learning pathways, including short courses, online platforms, and workplace-based learning. The presence of *badging* and *recognition* underscores the importance of credential visibility and acceptance, particularly in relation to employer recognition and career advancement. This aligns with broader trends towards unbundled and stackable learning models, where learners accumulate credentials across different platforms and contexts. However, this cluster appears more peripheral in the network and weakly connected to the central pedagogical clusters, suggesting that employability and recognition are not yet fully integrated into the dominant research discourse. This indicates a fragmentation between education-focused research and workforce-oriented applications, limiting the potential for micro-credentials to function as a cohesive ecosystem.

When considered together, Table 7, Figure 4, and Table 8 reveal a consistent structural pattern in which disciplinary foundations shape thematic development, which is then organised into distinct but interconnected research clusters. Table 8 confirms that these clusters are both structurally coherent and thematically meaningful. However, the distribution of clusters reveals uneven development across the field. The pedagogical domain remains the most dominant, while technological and employability-oriented research areas are still emerging and comparatively less integrated. When interpreted through the socio-technical-policy framework, the clusters reveal a multi-dimensional structure in which pedagogical, technological, and policy dimensions develop in parallel but remain only partially integrated. Pedagogical research, represented by Clusters 1 and 2, occupies a central position, emphasising competency-based learning, engagement, and assessment practices. In contrast, technological research (Cluster 3) reflects a growing but still evolving focus on digital credentialing systems, while policy and workforce-related research (Cluster 5) remains relatively peripheral and weakly connected to the core discourse. Cluster 4 plays a mediating role by linking pedagogical concepts with institutional implementation, highlighting the translation of theory into practice within educational ecosystems. To sum up, the five clusters indicate that micro-credential research is pedagogically anchored, with a strong emphasis on competency-based learning and engagement; technologically emerging yet not fully integrated with instructional practices; institutionally evolving, with increasing but fragmented implementation; and economically relevant, although still weakly connected to employability discourse. The co-existence of emerging terms such as blockchain and digital credentials alongside established educational concepts further indicates that the field is gradually transitioning towards implementation and system-level development. Nevertheless, the relatively limited integration between clusters indicates that these dimensions continue to develop in parallel rather than as a cohesive system. This fragmentation indicates that these dimensions are not yet fully integrated, limiting the development of a coherent micro-credential ecosystem. Consequently, advancing micro-credential research depends on bridging these domains through more integrative theoretical frameworks and empirically grounded studies that connect pedagogy, technology, and policy more systematically.

RQ4. Which publications have exerted the greatest intellectual influence on the field?

To examine intellectual influence comprehensively, this study integrates citation analysis (Table 9) with co-authorship network visualisation (Figure 5). While citation metrics identify highly influential publications, co-authorship patterns provide insight into how influence is produced and shared through collaborative research networks.

Table 9. Top 20 Highly Cited Articles

No.	Author(s)	Title	Source Title	Cites	Cites per Year
1	Moore et al. (2025)	Systematic review of digital microcredentials: trends in assessment and delivery	Distance Education	10	5
2	Bruguera et al. (2025)	Micro-credentials and soft skills in online education: the employers' perspective	Distance Education	7	3.5
3	Charles et al. (2024)	Micro-Credentials in Vocational and Professional Training	Integrating Micro-Credentials with AI in Open Education	6	2
4	Ironsi & Ironsi (2025)	Efficacy of micro-credential learning environments for developing students' 21st century skills: toward achieving sustainable development goals	International Journal of Educational Management	5	2.5

5	Gamage & Dehideniya (2025)	Unlocking Career Potential: How Micro-Credentials Are Revolutionising Higher Education and Lifelong Learning	Education Sciences	5	2.5
6	Durak & Çankaya (2024)	The Rise of Micro-Credentials: A New Certification System for Career Development	Integrating Micro-Credentials with AI in Open Education	5	1.666667
7	Datta et al. (2025)	Reskilling and Upskilling for Decarbonization: Analyzing Micro-Credential Programs for Energy Workforce Development	Journal of Continuing Higher Education	4	2
8	Charles & Hill (2025)	Micro-credentials and AI in higher education: Transforming professional development	Prompt Engineering and Generative AI Applications for Teaching and Learning	3	1.5
9	Hamutoğlu et al. (2024)	Triggering Competitiveness From a Global Perspective: MOOCs and Micro-Credentials	Integrating Micro-Credentials with AI in Open Education	3	1
10	Kovilpillai et al. (2024)	AI Enhanced Micro-Credentials for Efficiency and Accessibility: Using Gen-AI to Improve the Design, Development, and Delivery of Micro-Credentials	Integrating Micro-Credentials with AI in Open Education	3	1
11	Hou et al. (2025)	Does quality assurance matter for micro-credentials in higher education? Legitimacy, regulatory framework, and challenges in four Asian contexts	Studies in Higher Education	2	1
12	Varadarajan et al. (2025)	Institutional readiness for the implementation of micro-credentials in higher education	Distance Education	2	1
13	Patterson & Hepburn (2025)	Implementing a competency-based assessment approach to micro-credentials	Distance Education	2	1
14	Zhang & Alasmari (2025)	Micro-credentials in Saudi higher education: stakeholder perceptions and policy implications for economic transformation and bridging skills gaps	Journal of Professional Capital and Community	2	1
15	Sullivan (2025)	A Rubric for Microcredential Evaluation: Strengthening Quality Assurance	Innovative Higher Education	2	1
16	Tanim et al. (2025)	Dr. FANS: A Decentralized Micro-credential Verification System Towards Higher Qualifications	2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025	2	1
17	Stamatakis et al. (2024)	Micro-Credentials Establish New Realities in Digital Education	Integrating Micro-Credentials with AI in Open Education	2	0.666667
18	Halttunen et al. (2024)	Role of Micro-credentials and Open Badges in Sustainable Tourism Education	Nordic Coastal Tourism: Sustainability, Trends, Practices, and Opportunities	1	0.333333
19	LaRue Keeley (2025)	The Scope and Depth Micro-Credential Framework: Aligning	Innovative Approaches in Vocational and Regional Education	1	0.5

		Education With Evolving Industry Demands			
20	Narong (2025)	Research trends on micro-credentials: a keyword co-occurrence analysis and science mapping	Journal of Further and Higher Education	1	0.5

Source: Generated by the author(s) using biblioMagika® (Ahmi, 2024)

Table 9 indicates that intellectual influence in micro-credential research is concentrated in a relatively small number of recent publications. When citations are normalised using citations per year (C/Y), Moore et al. (2025) emerges as the most influential study (TC = 10; C/Y = 5), suggesting rapid scholarly uptake within a short period. Similarly, Bruguera et al. (2025) (C/Y = 3.5) and Ironsi and Ironsi (2025) (C/Y = 2.5) demonstrate strong citation performance, reflecting growing attention to employability, skills development, and learning outcomes. In addition, studies such as Gamage and Dehideniya (2025) and Hou et al. (2025) highlight the importance of institutional readiness and quality assurance, while Charles and Hill (2025) and Kovilpillai et al. (2024) reflect increasing interest in technological integration and digital credential systems.

These influential publications can be grouped into four thematic areas: (1) workforce alignment and employability, (2) institutional readiness and governance, (3) technological integration and digital infrastructure, and (4) sectoral and application-based contexts. The distribution of highly cited studies across these domains indicates that intellectual influence is not concentrated within a single research stream but is dispersed across multiple areas. From a Diffusion of Innovation (DOI) perspective, the prominence of recent publications with relatively high C/Y values suggests that the field is transitioning from early exploration to a growth phase, where new ideas are rapidly diffused but have not yet stabilised into a unified theoretical framework. From a Resource-Based View (RBV) perspective, the concentration of influential publications within certain journals and institutional contexts reflects the role of resource availability, research capacity, and network positioning in shaping scholarly impact.

No.	Theme	Example	Focus
1.	Workforce alignment and employability	Bruguera et al., 2025; Ironsi & Ironsi, 2025; Gamage & Dehideniya, 2025	Focusing on skills development and labour market relevance
2.	Institutional readiness and governance	Hou et al., 2025; Varadarajan et al., 2025; Patterson & Hepburn, 2025; Sullivan, 2025)	Addressing implementation, quality assurance, and policy alignment
3.	Technological integration and digital infrastructure	Charles & Hill, 2025; Kovilpillai et al., 2024; Durak & Çankaya, 2024	Focusing on platforms, AI, and digital credential systems
4.	Sectoral and application-based contexts	Datta et al., 2025; Tanim et al., 2025	linking micro-credentials to specific industries and contexts

To complement this citation-based perspective, the co-authorship network provides insight into how intellectual influence is structured through collaboration. Figure 5 shows that the network is organised into several colour-coded clusters, each representing a distinct research community. The yellow cluster, centred around Daniel T. Hickey and James E. Willis, reflects pedagogical and professional learning. The red cluster, including Kathryn S. Coleman, Sheryl Grant, and Kim Flintoff, represents digital infrastructure and credential systems. The green cluster, represented by Rupert Ward and Alan Hayes, captures learner engagement and skills development. The blue cluster, consisting of Jelena Jovanović, Bojan Tomić, and Vladan Devedžić, represents institutional and policy-oriented research.

Within these clusters, authors demonstrate strong intra-cluster collaboration, indicating cohesive and thematically aligned research communities. However, inter-cluster connections remain relatively weak, suggesting that collaboration is largely confined within specific domains. This structure indicates that knowledge production is organised around distinct thematic communities rather than a fully integrated network.

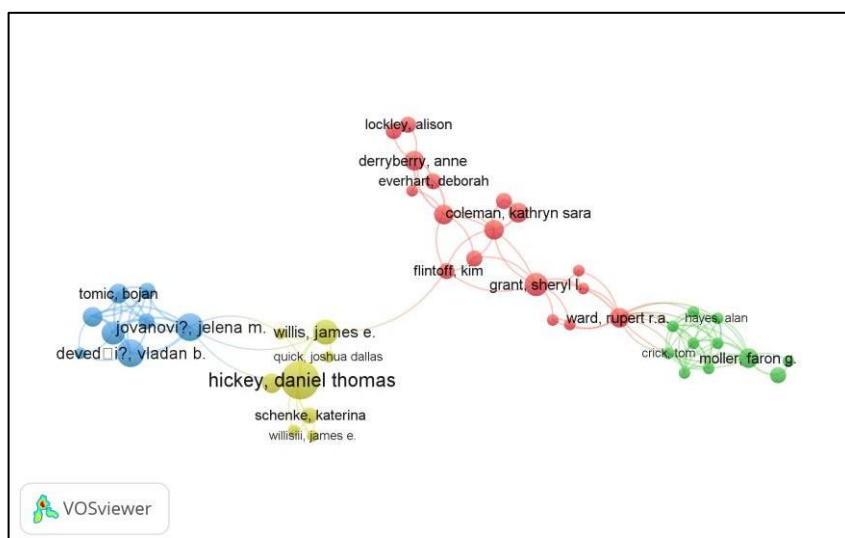


Figure 5. Network visualization of the author's co-authorship analysis

Source: Generated by the author(s) using VOSviewer (van Eck & Waltman, 2014)

Importantly, a comparison between Table 9 and Figure 5 reveals that the most highly cited publications are not necessarily produced by the most central authors within the co-authorship network. For example, highly cited studies such as Moore et al. (2025) and Bruguera et al. (2025) demonstrate strong citation impact but are not prominently positioned within the major collaboration clusters. Similarly, influential contributions on institutional readiness and governance, such as Gamage and Dehideniya (2025) and Hou et al. (2025), do not appear as central nodes within the co-authorship structure. This indicates that citation influence is distributed across multiple authors and thematic areas rather than concentrated within a single collaborative core.

Conversely, authors who are central within the co-authorship network, such as Hickey and Willis in the pedagogical cluster, and Jovanović and Devedžić in the institutional and policy-oriented cluster, demonstrate strong collaborative ties and contribute to cohesive research communities but do not feature prominently among the most highly cited publications in Table 9. This contrast indicates that authors who shape collaboration networks are not necessarily those whose individual publications receive the highest citation impact.

When the citation and co-authorship analyses are considered together, a clear structural pattern emerges. Highly cited publications contribute to the development of specific thematic domains but do not function as strong bridging points between clusters. As a result, intellectual influence is distributed across multiple areas rather than integrated within a unified research core. This confirms that influence in micro-credential research is not concentrated within a single group of highly collaborative scholars but is instead dispersed across distinct research communities.

Overall, the findings show that intellectual influence in micro-credential research is distributed rather than concentrated, emerging rather than fully established, and shaped through collaborative clusters rather than individual dominance. At the same time, the limited overlap between citation influence and collaboration structure highlights a fragmented intellectual landscape, where highly cited work and collaborative networks do not fully converge. This structural fragmentation suggests that the field

remains in a stage of consolidation, where theoretical coherence and interdisciplinary integration are still developing.

From a broader perspective, these findings indicate that micro-credential research is evolving as a multi-dimensional field, where pedagogical, technological, and policy-oriented domains develop at different rates and through distinct research communities. While the field is progressing towards greater maturity, its continued advancement depends on strengthening cross-cluster collaboration and integrating fragmented domains into a more cohesive and theoretically grounded framework.

5. Discussion

This study provides a comprehensive bibliometric analysis of micro-credential research between 2015 and 2025, addressing four research questions related to publication trends, research productivity, thematic development, and intellectual influence. Rather than restating descriptive findings, this discussion interprets the observed patterns by examining the underlying mechanisms that explain their emergence, the implications for the development of the field, and their contribution to existing knowledge.

The findings in RQ1 indicate that micro-credential research has evolved from a relatively niche area into a rapidly expanding and increasingly structured domain. Publication output increased steadily over the decade, with a marked acceleration after 2019 and a peak in 2024. This trajectory can be explained using the Diffusion of Innovation (DOI) framework, which posits that the adoption of innovations is shaped by factors such as relative advantage, compatibility, and observability (Rogers et al., 2014). Micro-credentials offer clear advantages in terms of flexibility, skills alignment, and responsiveness to labour market demands, making them attractive to higher education institutions (Gamage & Dehideniya, 2025; Varadarajan et al., 2023). Their compatibility with digital and online learning environments further supported rapid adoption, particularly during the COVID-19 pandemic, which accelerated the shift towards alternative credentialing models (Brown et al., 2023). In addition, increased visibility through institutional pilot initiatives and policy support enhanced observability, encouraging wider adoption across institutions (Selvaratnam & Sankey, 2021).

From a DOI perspective, the field appears to have progressed from early adoption towards the early majority stage, where micro-credentials are no longer experimental but increasingly embedded within institutional strategies (Ahsan et al., 2023; Tamoliūnė et al., 2023). This shift signifies entry into a consolidation phase. However, the relatively lower citation counts for recent publications, particularly in 2024 and 2025, reflect a citation maturation effect rather than declining relevance, as scholarly recognition typically lags behind publication output in emerging fields (Gomez et al., 2022; Nielsen & Andersen, 2021; Pruteanu et al., 2025). More importantly, the distribution of citations across numerous recent studies demonstrates that the intellectual foundations of the field are still forming, with influence dispersed rather than concentrated in a set of canonical works. This dispersion is further reinforced by the RQ4 findings, where highly cited publications are distributed across multiple thematic areas and are not concentrated within a single dominant research stream or collaborative core, indicating that intellectual consolidation is still in progress.

At the same time, the findings in RQ2 reveal clear disparities in research productivity and influence across countries, institutions, and authors. The United States dominates both publication output and citation impact, while countries such as Australia and Canada demonstrate relatively higher citation efficiency despite lower publication volumes (Brown et al., 2023; McGreal & Olcott, 2022; Varadarajan et al., 2023). This pattern can be interpreted through the Resource-Based View (RBV), which posits that systems with stronger resource configurations such as funding, digital infrastructure, research capacity, and policy support are more likely to achieve sustained competitive advantage (Barney, 1991). In addition, the dominance of English as the primary language of academic publishing further amplifies visibility and citation potential for institutions in the Global North (Arnado, 2021; Gomez et al., 2022).

These findings suggest that research influence is not solely determined by productivity but is also shaped by structural and institutional conditions, highlighting the need for more inclusive and globally distributed research collaboration. In particular, the divergence observed in RQ4 between citation influence and collaboration centrality suggests that access to resources not only affects publication output but also influences whether scholars are positioned within high-visibility citation networks or within cohesive collaborative communities.

The thematic analysis in RQ3 reveals that micro-credential research is organised into five major clusters: pedagogical foundations, learner engagement, technological infrastructure, instructional innovation, and employability pathways. While these domains are clearly identifiable, they tend to develop in parallel rather than as a fully integrated system. Pedagogical research remains dominant, focusing on competency-based learning, digital badges, and learner engagement (McGreal & Olcott, 2022; Pirkkalainen et al., 2022; Varadarajan et al., 2023). Technological research increasingly explores digital credentialing systems, including blockchain and platform-based infrastructures (Orman et al., 2023; Surono & Jusuf, 2025). Policy-oriented research addresses institutional readiness, employability, and quality assurance (Brown et al., 2023; Fonzo & Evangelista, 2025; Orman et al., 2023).

This thematic fragmentation is reinforced by the co-authorship network structure in RQ4, which reveals that collaboration is concentrated within clusters rather than across them. When interpreted together, the keyword and co-authorship analyses indicate that micro-credential research is not only conceptually segmented but also socially organised into relatively discrete research communities. Pedagogical clusters remain central and densely connected, while technological and policy-oriented clusters are comparatively peripheral and weakly integrated. This convergence between conceptual and collaborative fragmentation reflects that the field is developing through parallel trajectories rather than cumulative and interdisciplinary knowledge building. The RQ4 analysis further demonstrates that highly cited publications tend to contribute to specific clusters without acting as bridging mechanisms between them, reinforcing the persistence of structural fragmentation at both conceptual and collaborative levels.

From a socio-technical–policy perspective, this pattern reflects an imbalance in system development, where pedagogical innovation advances more rapidly than technological integration and policy alignment. According to DOI theory, such separation represents a transitional phase in the diffusion process, where different components of an innovation evolve independently before achieving full integration (Rogers et al., 2014). This indicates that the field has not yet reached conceptual or structural maturity, with diffusion operating across interconnected sub-systems rather than as a unified process.

The analysis of intellectual influence further supports this interpretation. Highly cited publications are relatively recent and distributed across multiple thematic areas, including workforce alignment, institutional readiness, technological innovation, and sector-specific applications. This distribution suggests that the field lacks a single dominant theoretical framework. Instead, influence is dispersed across multiple emerging domains, reflecting a field that is still forming its intellectual identity. Co-authorship analysis similarly shows that influence is generated through collaborative clusters rather than individual dominance, reinforcing the collective nature of knowledge production in this area. At the same time, the limited overlap between highly cited works and central collaboration nodes highlights that citation visibility and collaborative influence operate as distinct but interconnected mechanisms within the field.

In short, the cluster structure reveals that micro-credential research is pedagogically anchored, with strong emphasis on competency-based learning and learner engagement; technologically emerging, particularly in relation to digital credentialing systems; institutionally evolving, through growing but fragmented implementation practices; and economically relevant, although still weakly connected to employability discourse. This configuration indicates that the field remains structurally fragmented, with pedagogical, technological, and labour market dimensions developing in parallel rather than as an

integrated system. This structural separation may constrain the development of coherent theoretical frameworks and limit the capacity of research to inform holistic policy and institutional strategies.

To sum up, the findings confirm that micro-credential research is transitioning from early exploration towards structured consolidation. From a DOI perspective, the field is progressing towards broader institutional adoption, driven by perceived advantages and increasing visibility. From an RBV perspective, disparities in research influence reflect uneven resource distribution and institutional capacity. From a socio-technical–policy perspective, the field is multi-dimensional but not yet fully integrated. These findings contribute to the literature by demonstrating that micro-credential research should be understood as an interconnected socio-technical–policy system shaped by both innovation dynamics and resource configurations. While the field has expanded rapidly and gained global attention, its continued advancement depends on strengthening integration across pedagogical, technological, and policy domains. Future work should continue to explore mechanisms that enable stronger integration across these domains to support more coherent theoretical and practical development.

6. Implications and Future Research

The findings of this study generate important implications for theory, practice, and policy by providing a deeper understanding of how micro-credentials are developing as a complex and evolving field within higher education systems. Building on these findings, the following section outlines the implications for theory, practice, and policy.

6.1 Theoretical Implications

From a theoretical perspective, this study extends existing frameworks by demonstrating that the development of micro-credential research is not only driven by adoption dynamics but also shaped by structural and systemic imbalances. The observed publication growth and adoption patterns reinforce the relevance of Diffusion of Innovation (DOI), confirming that micro-credentials are transitioning from early experimentation towards broader institutional uptake (Rogers et al., 2014). However, the findings also reveal that diffusion does not occur uniformly across all dimensions of the innovation. Instead, pedagogical, technological, and policy components evolve at different rates, reflecting varying levels of complexity, compatibility, and institutional readiness. More importantly, the combined keyword and co-authorship analyses extend DOI by showing that diffusion in micro-credential research occurs unevenly across interconnected sub-systems rather than as a unified process. Pedagogical innovation appears to diffuse more rapidly, while technological and policy dimensions lag behind, indicating asynchronous adoption patterns. This suggests that DOI in complex educational innovations should be reconceptualised as a multi-layered and non-linear process, where different components of an innovation mature independently before achieving full integration. This reinforces the need to conceptualise diffusion as multi-layered and non-linear in complex educational systems.

At the same time, the Resource-Based View (RBV) provides a robust explanation for disparities in research productivity and influence. The dominance of countries such as the United States and the high citation efficiency of systems such as Australia reflect differences in institutional resources, including research funding, digital infrastructure, human capital, and policy support (Barney, 1991; Abdul Halim et al., 2024; Alsobhi et al., 2023). These findings indicate that research influence is not determined solely by output but by the strategic configuration and mobilisation of resources within institutional and national contexts.

This study advances the socio-technical–policy perspective by empirically demonstrating that fragmentation exists at both conceptual and collaborative levels. The alignment between keyword clusters and co-authorship clusters shows that micro-credential research is structured into relatively

discrete domains, where pedagogical, technological, and policy-related knowledge develops in parallel rather than through integrated frameworks. This underscores the need for future theoretical models that explicitly connect these dimensions and conceptualise micro-credentials as interconnected systems rather than isolated components (Ahsan et al., 2023; Selvaratnam & Sankey, 2021).

6.2 Practical Implications

From a practical perspective, the findings highlight the need for higher education institutions to adopt more coordinated and system-level approaches in implementing micro-credentials. The separation between pedagogical, technological, and policy domains suggests that institutional practices may also be fragmented. In many cases, instructional design, platform development, and quality assurance are handled independently, pointing to a system-level misalignment rather than isolated operational problems. Without coordinated integration, micro-credentials risk being implemented as fragmented initiatives that fail to deliver coherent learning pathways, recognised outcomes, or scalable solutions. Therefore, institutions should prioritise cross-functional governance models that align educators, instructional designers, IT specialists, and administrators within a unified framework.

To operationalise this integration, institutions may adopt a structured readiness approach that assesses key dimensions of implementation. These include alignment with national qualification frameworks, clarity of learning outcomes, robustness of assessment design, platform interoperability, employer relevance, and stackability into credit-bearing pathways. Practical indicators may include the proportion of micro-credentials mapped to qualification frameworks, employer uptake within a defined period, and the extent to which micro-credentials can be accumulated into formal qualifications, which together provide practical indicators for monitoring progress and effectiveness.

In addition, the increasing prominence of technological themes such as digital credentials and blockchain highlights the need for investment in robust digital infrastructure that supports secure verification, interoperability, and scalability (Alsobhi et al., 2023). However, technological development must be aligned with pedagogical principles to ensure that micro-credentials support meaningful learning outcomes rather than functioning solely as technical certifications. This requires integrating competency-based learning, assessment design, and learner engagement strategies within digital platforms (Tamoliūnė et al., 2023).

Furthermore, institutions in emerging regions, including Asia and the Middle East, can enhance their research visibility and impact by strengthening international collaboration and aligning micro-credential initiatives with recognised qualification frameworks. Such alignment improves credibility, supports global recognition, and enhances competitiveness within lifelong learning ecosystems (Ahsan et al., 2023; Varadarajan et al., 2023; Santally et al., 2024). Importantly, institutional strategies should focus not only on increasing publication output but also on improving research quality, collaboration networks, and interdisciplinary engagement.

6.3 Policy Implications

From a policy perspective, the study underscores the importance of developing coherent and integrated frameworks to support the recognition, portability, and scalability of micro-credentials. The fragmentation observed across thematic domains mirrors the cluster structure identified in the bibliometric analysis, where policy-related research remains peripheral and weakly connected to pedagogical and technological domains. This suggests that policy development has not kept pace with innovation in teaching and technology, highlighting a critical gap in governance alignment.

Policymakers should therefore prioritise the development of harmonised quality assurance frameworks that align national systems with international standards such as the European Qualifications Framework (EQF) and the ASEAN Qualifications Reference Framework (AQRf) (Oliver, 2019; Lang, 2023; McGreal et al., 2025). This is essential for ensuring the comparability, recognition, and transferability of micro-credentials across institutions and regions.

In addition, the uneven geographical distribution of research highlights the need for targeted capacity-building initiatives in underrepresented regions (Ahsan et al., 2023; Brown et al., 2021; Tamoliūnė et al., 2023; Selvaratnam & Sankey, 2021). Policymakers can support this by promoting international partnerships, funding collaborative research, and facilitating knowledge exchange between institutions with different levels of readiness (Gamage & Dehideniya, 2025; McGreal & Olcott, 2022). These efforts are essential to reduce global disparities and ensure that micro-credentials contribute to inclusive and equitable education systems.

Importantly, policy development should also align with global education priorities, particularly Sustainable Development Goal 4 (SDG 4), which emphasises inclusive access to quality education and lifelong learning opportunities (Ahad et al., 2025; Brown et al., 2021; McGreal et al., 2022; UNESCO, 2022; United Nations, 2023). By embedding micro-credentials within national education strategies, policymakers can support workforce development, enhance employability, and promote social mobility. However, achieving these outcomes requires strong alignment between governance structures, institutional capacity, and technological systems (Pirkkalainen et al., 2022; Varadarajan et al., 2023).

Overall, the implications of this study suggest that the future development of micro-credentials depends on the ability of institutions and policymakers to move beyond fragmented and parallel approaches towards integrated socio-technical-policy systems. Without such integration, the field risks remaining conceptually and operationally fragmented. By contrast, stronger alignment across pedagogy, technology, and policy can enhance the effectiveness, credibility, and scalability of micro-credentials, enabling them to function as a sustainable and transformative component of higher education and lifelong learning systems.

6.4 Future Research Directions

The findings of this study highlight several important gaps and emerging issues that provide clear directions for advancing micro-credential research. These directions are derived directly from the observed patterns in publication trends, thematic development, and intellectual structure, thereby offering a systematic pathway for future scholarship.

First, the rapid increase in publications after 2019 suggests that the expansion of micro-credential research cannot be explained solely by the COVID-19 pandemic. While the pandemic accelerated the adoption of online learning and flexible credentialing systems, the sustained growth beyond this period indicates the influence of broader institutional and policy drivers. Further studies should therefore examine the specific mechanisms that sustain adoption over time, including the role of national education policies, digital transformation strategies, and industry partnerships. From a Diffusion of Innovation (DOI) perspective, further studies could investigate how factors such as relative advantage, compatibility, complexity, and observability operate differently across institutional contexts and influence adoption trajectories at different stages of diffusion. Future work could also examine how these factors interact across pedagogical, technological, and policy dimensions to better understand variation in adoption patterns across contexts.

Second, the findings reveal significant disparities in research productivity and citation impact across countries and institutions. While some systems dominate in terms of output, others achieve higher citation efficiency with fewer publications. This raises important questions about the relationship

between research quality, visibility, and resource configurations. Future studies should explore how different combinations of resources such as funding, digital infrastructure, human capital, and international collaboration networks shape research influence. Drawing on the Resource-Based View (RBV), research could examine which resource configurations are most effective in generating high-impact scholarship and how institutions in underrepresented regions can strategically position themselves within global research networks.

Third, the thematic and co-authorship analyses reveal that micro-credential research is both conceptually and structurally fragmented, with pedagogical, technological, and policy domains developing in parallel rather than as an integrated system. Subsequent research should prioritise empirical investigations that examine how these domains can be effectively integrated in practice, rather than focusing solely on their conceptual separation. For example, empirical investigations could explore whether technological innovations such as blockchain and digital credential systems effectively support pedagogical goals, or whether pedagogical requirements drive technological adoption. In addition, studies are needed to examine how policy frameworks influence the alignment between curriculum design, assessment practices, and credential recognition. Addressing this gap is critical for moving the field towards a more coherent socio-technical–policy framework.

Fourth, although technological innovation is gaining increasing attention, particularly in areas such as blockchain, artificial intelligence, and digital credentialing systems, its relationship with educational outcomes remains under-theorised. Empirical research is needed to move beyond technical implementation and examine how these technologies influence learner engagement, assessment validity, and skill development. Future studies should focus on how digital credential systems support personalised learning pathways and how emerging technologies can be integrated with learner-centred pedagogical approaches, thereby helping to bridge the gap between technological advancement and educational effectiveness.

Fifth, the co-authorship network analysis reveals fragmented collaboration patterns, with strong intra-cluster connections but limited interaction across clusters. This suggests a need for research that examines the dynamics of interdisciplinary collaboration and its impact on knowledge development. Future studies could investigate whether more integrated collaboration networks lead to stronger theoretical contributions, more innovative methodologies, and greater practical relevance. In particular, research that brings together scholars from education, technology, and policy domains may provide valuable insights into how the field can overcome fragmentation and achieve greater conceptual coherence.

Sixth, the geographical distribution of research highlights the underrepresentation of regions such as Asia, Africa, and parts of the Middle East in high-impact publications. Although contributions from these regions are increasing, they remain less visible within global citation and collaboration networks. Future research should prioritise context-sensitive studies that examine how micro-credentials are implemented across diverse cultural, economic, and institutional environments. Ensuring inclusivity and representation of global diversity is critical to avoid the dominance of a limited number of regions.

Finally, the present study adopts a bibliometric approach and therefore provides a macro-level understanding of the field. Further studies should complement this perspective with empirical and longitudinal studies that examine the real-world impact of micro-credentials on learners, institutions, and labour markets. This includes investigating learner outcomes, employability trajectories, and the long-term effectiveness of micro-credential programmes in supporting lifelong learning. Longitudinal designs would be particularly valuable in capturing how micro-credentials evolve from pilot initiatives into fully integrated components of higher education systems.

Overall, these directions emphasise the importance of moving beyond isolated and descriptive approaches towards more integrated, empirically grounded, and context-sensitive research. In addition,

future research should prioritise integrative approaches that connect different domains of micro-credential research, enabling more cohesive theoretical development and practical implementation.

7. Limitations

Several limitations should be acknowledged when interpreting the findings of this study. These limitations are not only methodological but also shape the scope, representation, and generalisability of the conclusions drawn.

First, this study relies exclusively on the Scopus database. While Scopus provides extensive coverage of peer-reviewed literature and structured bibliographic data suitable for bibliometric analysis, it does not capture the full spectrum of micro-credential research. In particular, practitioner-oriented publications, policy reports, institutional documents, and grey literature are often excluded. This is a notable constraint given that micro-credentials are inherently practice-driven and closely linked to policy frameworks and institutional implementation. Consequently, the findings primarily reflect academic discourse rather than the broader ecosystem of practice and governance in which micro-credentials operate.

Second, the analysis is limited to English-language publications, introducing a linguistic bias that may systematically exclude relevant research produced in other languages. This is particularly significant for regions where micro-credentials are actively evolving, such as China, Latin America, and parts of Europe and Southeast Asia. As English dominates scholarly communication, this restriction may overrepresent Anglophone or internationally indexed systems while underrepresenting locally grounded perspectives. The combined use of Scopus and English-language filtering therefore biases the dataset towards globally visible scholarship, limiting the inclusivity of the analysis.

Third, citation-based indicators are subject to temporal and structural limitations. Recent publications, particularly those from 2024 and 2025, have had limited time to accumulate citations, which may underestimate their actual influence. Although normalisation techniques such as citations per year were applied, citation metrics inherently favour older publications. In addition, citation counts include self-citations, which may marginally inflate impact measures. These limitations should be considered when interpreting patterns of intellectual influence and scholarly prominence, particularly in emerging and rapidly evolving research areas.

Fourth, bibliometric analysis provides a macro-level perspective of research trends, thematic structures, and collaboration networks but does not capture the contextual depth of individual studies. While this approach is effective for identifying large-scale patterns, it cannot fully explain how micro-credentials are designed, implemented, or experienced within specific educational settings. As a result, the findings should be complemented by qualitative, mixed-method, and longitudinal studies that provide deeper insights into pedagogical practices, institutional dynamics, and learner outcomes.

Fifth, although systematic procedures were applied in data cleaning, keyword harmonisation, and clustering, a degree of subjectivity remains inherent in these processes. Decisions related to merging terms, standardising keywords, and interpreting thematic structures may influence the configuration of clusters and networks. While these steps were guided by established protocols and documented for transparency, they remain interpretive and may affect the representation of the field's intellectual structure.

To summarise, these limitations indicate that the findings represent a structured overview of English-language, Scopus-indexed scholarship rather than a fully comprehensive representation of global micro-credential research. Recognising these constraints is essential for appropriately contextualising the results and for guiding future research towards more inclusive, multi-source, and context-sensitive approaches.

8. Conclusion

This study set out to examine the development and structure of micro-credential research in higher education through a comprehensive bibliometric analysis of publications indexed in Scopus between 2015 and 2025. By integrating publication trends, productivity patterns, thematic clustering, and collaboration networks, the study provides a multi-dimensional understanding of how the field has evolved over time. The findings indicate that micro-credential research has grown rapidly, particularly after 2019, reflecting increasing institutional and policy interest in flexible and skills-based learning pathways. Despite this growth, the field remains unevenly developed, with significant disparities in research productivity and influence across countries, institutions, and authors. These disparities highlight the importance of resource availability, research capacity, and global visibility in shaping scholarly impact.

A key contribution of this study is demonstrating that micro-credential research is structured into multiple domains, including pedagogical design, learner engagement, technological infrastructure, instructional innovation, and employability pathways. These domains are reflected in distinct but only partially connected research communities, indicating that the field is developing through parallel research streams rather than as a fully integrated system. The analysis further demonstrates that intellectual influence is distributed across recent publications and collaborative clusters rather than concentrated within a small number of dominant authors or theoretical frameworks. This suggests that micro-credential research is still in a stage of consolidation, where its intellectual foundations are evolving and its conceptual boundaries are still being defined.

From a theoretical perspective, the study contributes to existing knowledge by extending Diffusion of Innovation (DOI) and Resource-Based View (RBV) frameworks. The findings show that adoption and development in micro-credential research occur unevenly across interconnected dimensions, reflecting a multi-layered diffusion process and resource-dependent patterns of influence. At the same time, the study advances a socio-technical-policy perspective by demonstrating that pedagogical, technological, and governance dimensions are interdependent but not yet fully aligned. From a practical and policy perspective, the findings highlight the need for greater integration across these domains to ensure that micro-credentials are implemented as coherent and scalable systems rather than fragmented initiatives. This includes aligning curriculum design, assessment practices, digital infrastructure, and policy frameworks to support effective and recognised learning pathways.

In conclusion, this study demonstrates that while micro-credential research has achieved significant growth and diversification, its continued advancement depends on strengthening integration across pedagogical, technological, and policy domains. Greater alignment across these dimensions will be essential for supporting coherent theoretical development, effective implementation, and meaningful contribution to higher education and lifelong learning systems.

ETHICAL APPROVAL

Not applicable

CONSENT TO PARTICIPATE

Not applicable

CONSENT TO PUBLISH

Not applicable

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this work the author(s) used ChatGPT in order to do the paraphrasing and some data interpretation. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

- Abdul Halim, F. S., Luaran, J. E., & Lee, S. S. J. (2024). Unravelling challenges of higher education institutions in implementing effective micro-credentials: A multi-stakeholder qualitative study. *Asian Journal of University Education (AJUE)*, 20(1), 114–126. <https://ir.uitm.edu.my/id/eprint/111479/>
- Ahad, M., Ahmed, F., Bhowmik, A., Mim, Z., Anu, A., & Busch, P. (2025). Exploring micro-credential affordances in Bangladeshi higher education. *Education and Information Technologies*, 30, 13565 - 13597. <https://doi.org/10.1007/s10639-025-13317-x>
- Ahsan, K., Akbar, S., Kam, B., & Abdulrahman, M. D. A. (2023). Implementation of micro-credentials in higher education: A systematic literature review. *Education and Information Technologies*, 28(10), 13505–13540. <https://link.springer.com/article/10.1007/s10639-023-11739-z>
- Ahmi, A. (2023). OpenRefine: An approachable tool for cleaning and harmonizing bibliographical data. *11th International Conference on Applied Science and Technology 2022 (11th ICAST 2022) AIP Conference Proceedings*, 2827, 030006-1-030006–030011. <https://doi.org/10.1063/5.0164724>
- Ahmi, A. (2024) biblioMagika, available from <https://bibliomagika.com>
- Alsobhi, H. A., Alakhtar, R. A., Ubaid, A., Hussain, O. K., & Hussain, F. K. (2023). Blockchain-based micro-credentialing system in higher education institutions: Systematic literature review. *Knowledge-Based Systems*, 265, 110238. <https://www.sciencedirect.com/science/article/abs/pii/S095070512201334X>
- Arnado, J. (2021). Structured inequalities and authors’ positionalities in academic publishing: The case of Philippine international migration scholarship. *Current Sociology*, 71, 356 - 378. <https://doi.org/10.1177/00113921211034900>
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.
- Brown, M., McGreal, R., & Peters, M. (2023). A strategic institutional response to micro-credentials: Key questions for educational leaders. *Journal of Interactive Media in Education*, 2023(1). <https://jime.open.ac.uk/articles/10.5334/jime.801>
- Brown, M., Mhichíl, M., Beirne, E., & Mac Lochlainn, C. (2021). The Global Micro-credential Landscape: Charting a New Credential Ecology for Lifelong Learning. *Journal of Learning for Development*. <https://doi.org/10.56059/jl4d.v8i2.525>
- Cudney, E., Anderson, S., Smith, A., Tran, S., Gouridine, C., Peterson, J., Gant, B., Stuhlman, S., Rousseau, B., Browning, A., & Wheeler, T. (2022). Micro-credentials in higher education: A systematic review of the literature. *The Transnational Journal of Business*. <https://doi.org/10.64010/emmt4350>

Fonzo, C., & Evangelista, L. (2025). The Role Of Quality Assurance In Supporting Micro-Credentials In European Education And Training Systems. *Community Notebook. People, Education and Welfare in the Society 5.0*. <https://doi.org/10.61007/qdc.2025.2.383>

Gamage, K. A. A., & Dehideniya, S. C. P. (2025). Unlocking Career Potential: How Micro-Credentials Are Revolutionising Higher Education and Lifelong Learning. *Education Sciences*, 15(5), 525. <https://doi.org/10.3390/educsci15050525>

Gomez, C., Herman, A., & Parigi, P. (2022). Leading countries in global science increasingly receive more citations than other countries doing similar research. *Nature Human Behaviour*, 6, 919 - 929. <https://doi.org/10.1038/s41562-022-01351-5>

Guidolin, M., & Manfredi, P. (2023). Innovation diffusion processes: Concepts, models, and predictions. *Annual Review of Statistics and Its Application*, 10, 451–473. <https://doi.org/10.1146/annurev-statistics-040220-091526>

Guo, Q., & Huang, W. (2024). Analyzing the diffusion of innovations theory. *Scientific and Social Research*, 6(12). <https://doi.org/10.26689/ssr.v6i12.8947>

Harizan, S. H. M., & Ally, M. (2025). Artificial intelligence in micro-credentials for open and distance learning: A technologically-enhanced systematic review. In *Integrating Micro-Credentials With AI in Open Education* (pp. 238-278). IGI Global. <https://doi.org/10.4018/979-8-3693-5488-9.ch014>

Hou, A. Y. C., Tao, C. H. Y., Su, E. H. C., Zhou, K. Z. W., Chen, Y., Lin, A. F. Y., & Hill, C. (2025). Does quality assurance matter for micro-credentials in higher education? Legitimacy, regulatory framework, and challenges in four Asian contexts. *Studies in Higher Education*, 1–17. <https://www.tandfonline.com/doi/abs/10.1080/03075079.2025.2482833>

Irwanto, I., Wahyudiati, D., Saputro, A. D., & Lukman, I. R. (2023). Massive open online courses (MOOCs) in higher education: A bibliometric analysis (2012–2022). *IJIET: International Journal of Information and Education Technology*, 13(2), 223–231.

Johan, R. C., Rullyana, G., Susilana, R., & Emilzoli, M. (2025). Micro-Credentials in Higher Education: A Review and Bibliometric. *Journal of Education and Learning (EduLearn)*, 19(2), 1103–1116. <https://eric.ed.gov/?id=EJ1464221>

Krishnan, K., Joshith, V. P., Chittakath, S., & Ranjan, S. (2025). Trends and trajectories in MOOCs research terrain: A bibliometric mapping from 2008 to 2023. *Journal of Education and Learning (EduLearn)*, 19(4), 2266–2278. <https://edulearn.intelektual.org/index.php/EduLearn/article/view/23268>

Lang, J. (2023). Workforce upskilling: can universities meet the challenges of lifelong learning?. *The International Journal of Information and Learning Technology*, 40(5), 388-400. <https://www.emerald.com/ijilt/article-abstract/40/5/388/134512/Workforce-upskilling-can-universities-meet-the?redirectedFrom=fulltext>

- Martín-Ramos, P., Diaz-De-Oleo, D., Fourati-Jamoussi, F., Couchy, K., Lo Giudice, L., Tosi, B., Oliveira-Pinto, F., Veiga-Martins, L., & Navas-Gracia, L. (2025). Digital credentials for environmental competencies: A scoping review of open recognition systems in climate education. *Open Research Europe*, 5. <https://doi.org/10.12688/openreseurope.20048.3>
- McGreal, R., Mackintosh, W., Cox, G., & Olcott Jr, D. (2022). Bridging the gap: Micro-credentials for development: UNESCO chairs policy brief form-under the III world higher education conference (WHEC 2021) type: Collective X. *International Review of Research in Open and Distributed Learning*, 23(3), 288-302.
- McGreal, R., & Olcott Jr, D. (2022). A strategic reset: Micro-credentials for higher education leaders. *Smart Learning Environments*, 9(1), 9. <https://link.springer.com/article/10.1186/s40561-022-00190-1>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Bmj*, 339.
- Moodie, G., & Wheelahan, L. (2021). Credentialing micro credentials. *Journal of Teaching and Learning for Graduate Employability*, 12(1), 58–71. <https://ojs.deakin.edu.au/index.php/jtlge/article/>
- Narong, D. K. (2025). Research trends on micro-credentials: A keyword co-occurrence analysis and science mapping. *Journal of Further and Higher Education*, 49(7), 919-937. <https://doi.org/10.1080/0309877X.2025.2509578>
- Ngoc Ha, N. T., Spittle, M., Watt, A., & Van Dyke, N. (2023). A systematic literature review of micro-credentials in higher education: A non-zero-sum game. *Higher Education Research & Development*, 42(6), 1527–1548. <https://www.tandfonline.com/doi/abs/10.1080/07294360.2022.2146061>
- Ngoc Ha, N. T., Van Dyke, N., & Spittle, M. (2025). Micro-credentials in higher education: perceived benefits for graduate employability and interest levels in micro-credentials for training employability skills. *Studies in Higher Education*, 1–13. <https://www.tandfonline.com/doi/full/10.1080/03075079.2025.2516709>
- Nielsen, M. W., & Andersen, J. P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7), e2012208118. <https://doi.org/10.1073/pnas.2012208118>
- Oliver, B. (2019). *Making micro-credentials work for learners, employers and providers* (pp. 1435129097–1637135585). chrome-extension://efaidnbmnnnibpcajpcgglefindmkaj/https://ltr.edu.au/resources/Making-micro-credentials-work-Oliver-Deakin-2019-full-report.pdf
- Orman, R., Şimşek, E., & Kozak Çakır, M. A. (2023). Micro-credentials and reflections on higher education. *Higher Education Evaluation and Development*, 17(2), 96–112. <https://doi.org/10.1108/HEED-08-2022-0028>

- Pirkkalainen, H., Sood, I., Padron Napoles, C., Kukkonen, A., & Camilleri, A. (2023). How might micro-credentials influence institutions and empower learners in higher education? *Educational Research*, 65(1), 40–63. <https://doi.org/10.1080/00131881.2022.2157302>
- Pruteanu, M. V., Moroşanu, A., Zegan, G., et al. (2025). Highlighting global inequities in health services quality research: A systematic review and quantitative evidence (2014–2023). *Health Research Policy and Systems*, 23, 105. <https://doi.org/10.1186/s12961-025-01376-y>
- Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432–448). Routledge.
- Rogers, G., Szomszor, M., & Adams, J. (2020). Sample size in bibliometric analysis. *Scientometrics*, 125, 777–794. <https://doi.org/10.1007/s11192-020-03647-7>
- Rullyana, G., Siregar, E., & Kustandi, C. (2025). Research trends on micro-credentials in higher education: A bibliometric analysis using Scopus and WoS databases. *Journal of Learning for Development*, 12(3), 484–500. <https://doi.org/10.56059/jl4d.v12i3.1950>
- Santally, M. I., Rajabalee, Y. B., Cooshna-Naik, D., & Greller, W. (2024). Perception of Education and Industry Leaders on Micro-Credentials and Their Potential in Higher Education. *Journal of Learning for Development*, 11(3), 528–539. <https://eric.ed.gov/?id=EJ1449740>
- Šarcevic, I. (2024). A New Agenda on Micro-Credentials: Filling the Gaps in the European Approach. *Journal of Learning for Development*, 11(1), 181–186. <https://eric.ed.gov/?id=EJ1423598>
- Savita, K. S., Isawasan, P., Ahmad Asmawi, M. A. H., Shaheen, M., & Ghafoor, R. (2025). Emerging themes and research directions in MOOCs and micro-credentials. *International Journal of Modern Education and Computer Science*, 17(6), 97–110. <https://doi.org/10.5815/ijmecs.2025.06.07>
- Selvaratnam, R., & Sankey, M. (2021). The State of Micro-Credentials Implementation and Practice in Australasian Higher Education. *Open Praxis*, 13. <https://doi.org/10.5944/openpraxis.13.2.130>
- Selvaratnam, R. M., & Sankey, M. D. (2021). An integrative literature review of the implementation of micro-credentials in higher education: Implications for practice in Australasia. *Journal of Teaching and Learning for Graduate Employability*, 12(1), 1–17. <https://search.informit.org/doi/abs/10.3316/informit.961591765882718>
- Strielkowski, W. (2025). The role of micro-credentials in the future digitalized AI-driven education. *arXiv preprint arXiv:2511.07583*.
- Surono, S., & Jusuf, H. (2025). Future-proofing graduates: A dynamic micro-credential quality management system in higher education to ensure self-learning ability, innovation capacity, and career growth. *Edunity*, 4(3). <https://doi.org/10.57096/edunity.v4i3.385>
- Takahashi, C. K., de Figueiredo, J. C. B., & Scornavacca, E. (2024). Investigating the diffusion of innovation: A comprehensive study of successive diffusion processes through analysis of search trends, patent records, and academic publications. *Technological Forecasting and Social Change*, 198, 122991.
- Tamoliūnė, G., Greenspon, R., Tereseviciene, M., Volungeviciene, A., Trepule, E., & Dauksiene, E. (2023, January). Exploring the potential of micro-credentials: A systematic literature review. *Frontiers*

in *Education*. 7,
1006811. <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2022.1006811/full>

UNESCO. (2022). *Towards a common definition of micro-credentials*. UNESCO Digital Library. <https://unesdoc.unesco.org/ark:/48223/pf0000381668>

United Nations. (2023). *Sustainable Development Goal 4: Quality Education*. <https://www.un.org/sustainabledevelopment/education/>

Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact: Methods and practice* (pp. 285–320). Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-10377-8_13

Varadarajan, S., Koh, J. H. L., & Daniel, B. K. (2023). A systematic review of the opportunities and challenges of micro-credentials for multiple stakeholders: learners, employers, higher education institutions and government. *International Journal of Educational Technology in Higher Education*, 20(1), 13. <https://link.springer.com/article/10.1186/s41239-023-00381-x>

Varadarajan, S., Koh, J. H. L., & Daniel, B. K. (2025). Institutional readiness for the implementation of micro-credentials in higher education. *Distance Education*, 46(1), 36–55. <https://www.tandfonline.com/doi/abs/10.1080/01587919.2024.2442017>

Vargo, S. L., Akaka, M. A., & Wieland, H. (2020). Rethinking the process of diffusion in innovation: A service-ecosystems and institutional perspective. *Journal of Business Research*, 116, 526–534. <https://doi.org/10.1016/j.jbusres.2020.01.038>

Vordenberg, S. E., Fusco, N. M., Ward, K. E., Darley, A., Brady, J. H., Culhane, N. S., Habib, M. J., Hernandez, E., Moye, P. M., Munusamy, S., Painter, J. T., Pope, N., Stevenson, T. L., Vanderboll, K., Chase, P. A., & Matsumoto, R. R. (2024). An integrative review of micro-credentials and digital badges for pharmacy educators. *American Journal of Pharmaceutical Education*, 88(3), 100660. <https://doi.org/10.1016/j.ajpe.2024.100660>

Yüksel, A. O. (2025). Investigation of studies on micro-credentials by bibliometric analysis method. *Kastamonu Education Journal*, 33(4), 753–767. <https://doi.org/10.24106/kefdergi.1795795>

Zakaria, R., Ahmi, A., Ahmad, A. H., Othman, Z., Azman, K. F., Ab Aziz, C. B., ... & Shafin, N. (2021). Visualising and mapping a decade of literature on honey research: a bibliometric analysis from 2011 to 2020. *Journal of Apicultural Research*, 60(3), 359–368.

Zhang, M., Thien, L. M., & Ahmi, A. (2025). Profiling the research landscape of student leadership from 1930 to 2023: A bibliometric analysis. *European Journal of Education*, 60, e70043. <https://doi.org/10.1111/ejed.70043>

Use of Artificial Intelligence-Based Writing (AIW) Tools in Higher Education: Opportunities, Challenges, and Future Directions for Sustainable Academic Practices

ABSTRACT

Authors: Lilliaty Ismail^{1*}, Sahar M. Taresh¹, Nooreen Noordin¹, Ilyana Jalaluddin²

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

²Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

*Corresponding author:

liliati@upm.edu.my

Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

The use of AI-based writing (AIW) tools in academic writing has become increasingly common in higher education. At the same time, their implications for writing pedagogy and academic integrity remain widely debated. This systematic review synthesises research published between January 2020 and August 2025 on the benefits and challenges of AIW tools in higher education, with particular attention to pedagogy, ethics, inclusivity, and institutional policies. Peer-reviewed studies were identified through searches of major databases, including Scopus, Web of Science (Core Collection), and Embase. Studies were screened and selected using predefined inclusion and exclusion criteria. Due to variation in study designs and outcome measures, the findings were synthesised using a narrative thematic approach. Methodological quality and potential sources of bias were assessed using the Mixed Methods Appraisal Tool (MMAT, 2018), and study quality was considered when interpreting the results. A total of 27 studies were included in the review. The findings show that AIW tools can support academic writing by providing immediate feedback, assisting non-native English speakers, and improving writing efficiency. However, the review also identifies significant concerns related to academic integrity, authorship attribution, equitable access, and possible negative effects on higher-order thinking skills. The existing evidence base is limited by short-term study designs, context-specific findings, and a lack of longitudinal and policy-focused evaluations. Overall, this review highlights the need for responsible, fair, and sustainable integration of AIW tools in higher education. Such integration should balance technological opportunities with core educational values and contribute to the goal of Quality Education (SDG 4). This systematic review was registered with the Open Science Framework (OSF) (10.17605/OSF.IO/F8P6C)

Keywords: Artificial intelligence (AI), higher education, academic writing, academic integrity, AI-based writing (AIW) tools.

1. Introduction

It is undeniable that artificial intelligence (AI) has significantly changed the educational landscape at all levels including in higher education. This includes the use of AI-based writing (AIW) tools in academic writing. AIW tools include Grammarly, ChatGPT, WriteFull, and QuillBot. The use of these tools have significantly transformed the manner in which students engage with their academic writing tasks (Imran & Norah, 2023; Moorhouse, Wan, Wu, Wu, & Ho, 2025).

Academic writing is an integral part of the higher education experience. There are rules and conventions in academic writing which students must master. These include presenting coherent and well-substantiated discussions and arguments, using appropriate academic English and adhering to citation regulations (Hind, 2024). Synthesizing, analyzing and evaluating a huge amount of past literature for the literature review, identifying gaps from past literature and discussing how the present study addresses these gaps can be a highly demanding task for students. In addition, they would also need to identify and discuss theories in relation to their study. All these tasks are cognitively demanding and time consuming. For many students, the challenges in academic writing are further compounded by lack of English language proficiency. Thus, AIW tools are viewed as a potential solution to overcoming these challenges.

However, the opportunities afforded by AIW tools come together with great ethical concerns. The ease and efficiency in which they generate sophisticated and linguistically coherent content, has triggered widespread concern of academic misconduct and decline in critical thinking and writing mastery among students. The lack of intellectual contribution, originality and critical engagement in the writing is particularly concerning as the assessment and grade the student receives might not be a true reflection of the students' ability and content mastery (e.g., Hazem, John, & Mike, 2023; Johnston, 2024). Critics also warn that excessive reliance on AI may undermine the development of critical thinking, analytical reasoning, and independent writing skills competencies that are central to the purpose of higher education (Gallent, Zapata, & Ortego, 2023; Kleopatra, 2024).

Higher education institutions globally have started to respond to these downsides of AIW tools by developing suitable guidelines and policies to ensure responsible use of AIW tools. However, these institutions face challenges in determining the permissible level for using AI, and the effectiveness and credibility of AI-detection tools. Often AI-detection tools cannot make definite determination of AI use and can only suggest likelihood of AI use. The rapid and dynamic revolution of AI also makes developing guidelines a challenge for higher education institutions. As AI continues to grow in robustness and sophistication, guidelines can quickly become obsolete. In addition to pedagogical and ethical considerations, the sustainable integration of AIW tools, aligned with the goal of Quality Education (SDG 4) has become an increasingly important concern for higher education institutions.

Recent systematic reviews have also started studying the use of artificial intelligence writing (AIW) tools in higher education; although the scope is still comparatively limited and divided. For example, recent reviews have concentrated more on general AI uses in education or initial evidence about the tools like ChatGPT, and they tend to be more tuned to the short-term pedagogical effects and ethical considerations without considering the bigger institutional picture (e.g., Kasneci et al., 2023; Zawacki-Richter et al., 2019). Although these studies are extremely valuable as frameworks, they are limited in the synthesis that would integrate governance, inclusiveness, and sustainability as an analytical framework. In addition, the reviews before 2022 can be dated by the time the new generation AI technologies got out of the rapid growth phase, which restricts their applicability to the present educational setting. On the contrary, the current systematic review builds on the current body of knowledge in three major aspects. First, it covers a period that is more recent and urgent (2020-2025), both the digital transformation driven by the pandemic and the advent of more powerful generative AI systems. Second, it uses a multidimensional prism that clearly incorporates the elements of pedagogical, ethical, governance, inclusivity, and sustainability, thus filling gaps in previous syntheses. Third, it improves methodological clarity by registering on Open Science Framework (OSF) and providing

extensive supporting material, which facilitates reproducibility and rigor in accordance with modern standards of systematic review.

To bring even more analytical focus on the review, the objectives are operationalized by a set of guiding sub-questions. First, in what ways do various categories of AI-based writing aids (i.e. grammar and style aids, paraphrasing aids, and generative AI systems (i.e. ChatGPT)) differ in terms of pedagogical value and risks? Second, how do the stated advantages and difficulties differ between the student groups, specifically between multilingual (L2) student learners and native (L1) speakers, and learners who have a low level of academic achievements and those who have a high level of academic achievements? Third, which institutional and contextual conditions mediate the effective and responsible use of AIW tools in higher education including assessment design, policy frameworks, and access conditions? These sub-questions enhance the overall review goals since they allow differentiated and theoretically informed synthesis of the literature especially regarding tool functionality, learner attributes and institutional governance.

2. Theoretical and Conceptual Framework

This systematic review is grounded in well-established theories of learning and writing. In particular, it draws on the sociocultural theory and self-regulated learning theory, both of which have long been used to explain how learners develop complex skills such as academic writing.

From a sociocultural perspective, learning takes place through mediated support that is gradually reduced as learners gain competence (Vygotsky, 1978; Wood, Bruner, & Ross, 1976). In writing research, scaffolding commonly refers to instructional supports such as feedback, modelling, prompts, and guided revision that help students move from novice to more independent writers. More recent research in educational technology has extended this idea by viewing digital tools as cognitive or pedagogical scaffolds that support different stages of the writing process (Graham & Harris, 2018; Ranalli, 2018).

Alongside this, self-regulated learning theory emphasises the active role of learners in planning, monitoring, and evaluating their own learning (Zimmerman, 2002). Academic writing is widely recognised as a self-regulated activity that involves drafting, revising, evaluating sources, and responding to feedback. Technologies that provide support during drafting and revision may therefore influence not only the quality of written texts but also how learners regulate their writing processes.

Building on these perspectives, this review adopts a conceptual framework that views AI-based writing (AIW) tools as scaffolding supports within the writing process. Rather than treating AIW tools as replacements for learning, the framework conceptualises them as supports that shape key learner processes, including drafting, revising, evaluating content, and reflecting on writing decisions. This view is supported by several empirical studies included in the review, which show that students commonly use AIW tools for early drafting, feedback, and revision support (e.g., Angelos et al., 2024; Oates & Johnson, 2025; King et al., 2024).

At the same time, drawing on work in academic integrity and educational ethics, the framework recognises that scaffolding can become counterproductive when it is poorly guided or misused (Bretag, 2016; Eaton, 2021). When AIW tools are used without clear pedagogical structure, students may bypass cognitive effort, leading to concerns related to authorship, plagiarism, and surface-level learning. These risks are documented in several of the reviewed studies (AlAfnan et al., 2023; Duah & McGivern, 2024).

The framework also incorporates institutional governance as a key moderating layer. Research on educational policy and assessment design shows that learning technologies do not operate in isolation; their effects are shaped by institutional policies, assessment practices, lecturer guidance, and access conditions (Bearman et al., 2022; Kővári, 2025). Consistent with this view, the reviewed studies suggest

that factors such as institutional policies, assessment redesign, staff training, and equity-focused infrastructure strongly influence whether AIW tools function as productive learning supports or as sources of ethical and educational risk.

Figure 1 presents the conceptual framework adopted in this review. The framework links AIW tool affordances to learner processes and learning outcomes, while highlighting the central role of institutional governance in shaping responsible, ethical, and sustainable use aligned with Quality Education (SDG 4).

Drawing on scaffolding theory and self-regulated learning, the framework positions AIW tools as pedagogical supports that influence proximal writing processes, including drafting, revision, evaluation, and reflection. These processes, in turn, shape more distal outcomes such as writing skills, critical thinking, and academic integrity. Institutional governance mechanisms such as policy development, assessment design, lecturer training, access, and sustainability considerations moderate these relationships and guide responsible AI integration. This framework informed the data extraction and thematic analysis by guiding the categorisation of pedagogical, ethical, inclusivity, and governance-related findings.

Finally, research on AIW tools has evolved rapidly since 2020, particularly following the introduction of large language models such as ChatGPT. By adopting a PRISMA-based systematic review approach, this study aims to provide educators, institutional leaders, and policymakers with a clear, evidence-based understanding of the opportunities and challenges that AIW tools present for higher education.

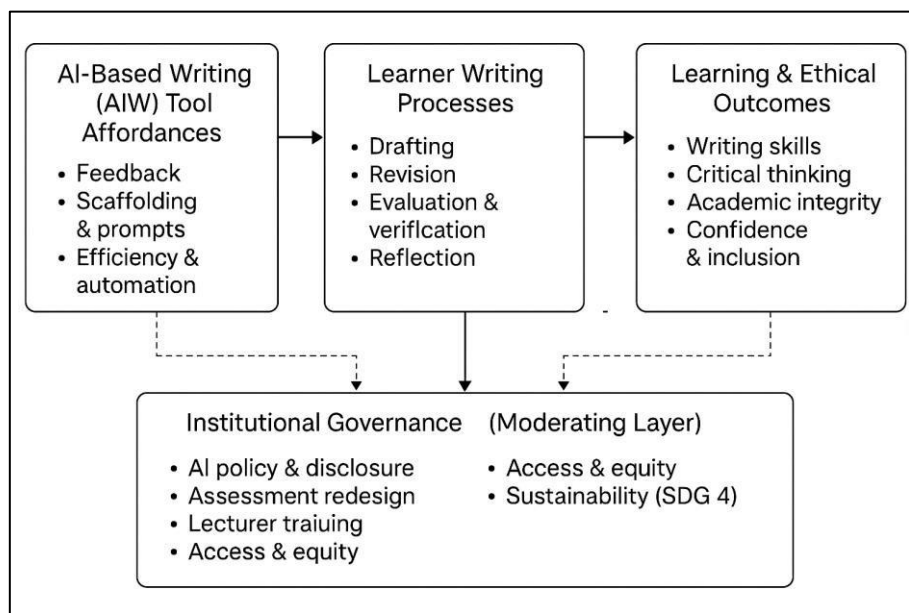


Figure 1. Processes, Learning Outcomes, and Institutional Governance

3. Methodology

3.1. Review Design

The current study was conducted in two phases. The first phase involved a systematic review, followed by a thematic analysis. The recommendations of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) statement were followed for conducting the systematic literature review. The following subsections detail the search strategies, eligibility criteria, data extraction, quality assessment, and analytical methods.

For the review, publications were collaboratively collected and evaluated by the first and second researchers using Google Drive, which facilitated the planning and management of the systematic literature review. This study was conducted using the systematic literature review methodology because it provides a comprehensive and clear overview of the literature written to date and identifies gaps and future research topics.

3.2. Review Process

The current review followed the five-step methodology for systematic literature reviews proposed by Denner and Tranfield (2009), which divides the review process into question formulation, study identification, study selection and evaluation, analysis and synthesis, and results reporting. The review protocol was retrospectively registered on the Open Science Framework (OSF) platform after the initial search and screening process was completed (OSF; DOI: 10.17605/OSF.IO/F8P6C). Such a retrospective registration was conducted to increase transparency and make sure that the review processes would be documented publicly. Notably, the research questions, inclusion criteria, search strategy, and data extraction procedures were not substantially modified once the screening process started. The registered protocol thus clearly shows the methodological approach applied in this study. Moreover, the OSF platform offers a record of the protocol under version control, which allows complete transparency about the moment of registration and the content. In addition, the scopes, eligibility criteria, and general analytical constructs of the study were not influenced by any minor procedural clarifications that were made during the review process.

The overall review flow also aligns with the PRISMA 2020 guidelines, and the main stages are summarized in Figure 2, with details of each stage provided.

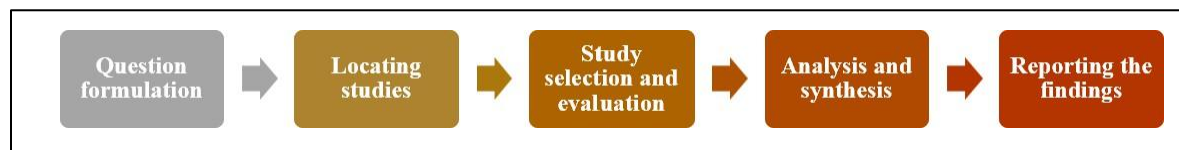


Figure 2. Steps in Conducting the Systematic Literature review (adapted from Denyer and Tranfield, 2009)

3.2.1. Formulating the Questions

The first step in the systematic review process is formulating the questions. The review's focus was defined as the role of AI-powered writing tools in higher education. After several rounds of discussion among the researchers, four guiding questions were agreed upon to examine the following: (1) What AI-powered writing tools are used in university settings? (2) How do these tools affect academic writing and learning processes? (3) What challenges related to ethics, equity, and integrity are present? (4) What governance and sustainability issues have been discussed in the literature?

These questions were developed according to a population-concept-context logic, where higher education represents the population/context, and AI-powered writing tools are the central concept. These questions were used to guide and systematically develop the research strategy, eligibility criteria, and data extraction model structure, ensuring alignment between the review objectives, data collection, and objective analysis. Each research question was addressed through thematic grouping, with topics explicitly linked to guiding questions to ensure analytical transparency (see Supplementary File A).

3.2.2. Study Identification

The second step was locating studies, involved designing and implementing a comprehensive search strategy. We searched three electronic databases Scopus, Web of Science (Core Collection) (full), and Embase, because together they cover education, psychology, health, and interdisciplinary higher education research. The final searches were conducted between January and August 2025 (UTC).

A pilot search was first run in Scopus to test and refine keywords and controlled vocabulary. The final search strategy combined terms related to artificial intelligence and generative AI (e.g., “*artificial intelligence*,” “*AI*,” “*ChatGPT*,” “*GPT-4*,” “*generative AI*,” “*AI writing tools*”), *academic writing* (e.g., “*academic writing*,” “*writing*,” “*composition*”), and *higher education* (e.g., “*higher education*,” “*university*,” “*college*,” “*postsecondary*”), using Boolean operators AND and OR. Search limits were applied to restrict results to studies published between 2020 and 2025, written in English, and indexed as peer-reviewed articles or conference proceedings. Similar strategies were adapted for each database to account for differences in indexing and syntax. The complete search strategies for all databases, including exact Boolean strings and field tags, are provided in Supplementary File B, reported in accordance with PRISMA-S guidelines. All retrieved references were exported to EndNote for initial management and removal of duplicate records, and the reference lists of included studies were manually screened to identify additional relevant publications. No additional grey literature sources were searched.

3.2.3. Study selection and evaluation

The third step, study selection and evaluation, focused on screening the located records against predefined inclusion and exclusion criteria derived from the review questions. We included studies published in peer-reviewed journals or conference proceedings between January 2020 and August 2025, written in English, and situated in higher education contexts (undergraduate, postgraduate, or academic staff). To be eligible, studies had to examine the use of AI-based tools for academic writing, such as generative AI chatbots, grammar and style assistants, or paraphrasing tools, and report on at least one of the following aspects: pedagogical or learning outcomes, academic integrity or ethical issues, equity and accessibility, or institutional policies and governance. We excluded publications that focused on primary or secondary schooling, discussed AI in education without a clear link to writing, were opinion pieces, editorials, or news items without an empirical or clearly described methodological basis, were not available in full text, or were written in languages other than English.

After deduplication in EndNote, all remaining records were imported into Google drive (Excel) to facilitate independent and blinded screening by two reviewers. Titles and abstracts were first screened to remove clearly irrelevant studies; full texts were then retrieved for the remaining records and assessed against the same eligibility criteria in Google Drive (Excel). Disagreements at both stages were resolved through discussion and consensus. At this stage, we (first and corresponding authors) also evaluated the methodological quality of each included study using the Mixed Methods Appraisal Tool (MMAT, 2018). Quality appraisal was conducted independently by two reviewers, focusing on study design, clarity of research questions, sampling procedures, and appropriateness of analysis. Quality ratings were not used as exclusion criteria but were used to inform the interpretation and weighting of evidence during synthesis. A summary of MMAT assessments for all included studies is provided in Supplementary File C. The study selection process is illustrated in PRISMA Figure 3.

To limit the search strategy to peer-reviewed journal articles, methodological rigor, quality, and consistency of the evidence base were ensured. Grey literature sources (e.g. institutional reports, policy documents, conference proceedings, theses) were excluded, as such sources may differ in terms of methodological transparency and are not always indexed in databases like Scopus, Web of Science and Embase. The reason behind this choice is that the study is aimed at integrating empirically validated results into a regulated and comparable data set.

To enhance interpretability of the findings, the strength of evidence supporting each theme was qualitatively assessed based on the number, methodological quality, and consistency of the included studies. Evidence strength was categorized into three levels:

1. **Strong evidence:** supported by multiple studies (≥ 5) with consistent findings, including at least two studies of high methodological quality based on MMAT criteria

2. **Moderate evidence:** supported by a moderate number of studies (2–4), with generally consistent findings but some variation in quality or results
3. **Limited evidence:** supported by one study or a small number of studies with inconsistent findings or lower methodological quality

This approach allowed for a more comprehensive interpretation of the results and facilitated transparency in linking conclusions to the underlying evidence base. A detailed breakdown of exclusion reasons at both the title/abstract and full-text screening stages is provided in Table 1, and is fully consistent with the PRISMA flow diagram (Figure 3), ensuring transparency and reproducibility.

3.2.4. Analysis and Synthesis

In the fourth stage, analysis and synthesis, data from the included studies were extracted into a structured template. For each study, bibliographic information, country, research design, participants, type of AI-based writing (AIW) tool, educational context, and key findings related to pedagogical outcomes, academic integrity, inclusivity, and institutional governance were recorded.

A thematic synthesis approach was employed, combining deductive and inductive coding strategies. Initial coding categories were informed by the study objectives and conceptual framework, while additional themes emerged inductively from the data. The unit of analysis was the key findings reported in each included study. The extracted data were read repeatedly to identify recurring patterns.

Two reviewers independently conducted initial open coding using a structured coding framework (see Supplementary File E). Coding and data management were supported using Microsoft Excel. An initial calibration exercise was conducted on a subset of studies to align coding interpretations. Discrepancies between reviewers were resolved through iterative discussion and consensus.

With the help of an iterative process of comparison and refinement, the codes were clustered into descriptive sub-themes, and then grouped into four overarching analytical themes, namely pedagogical benefits, ethical and integrity challenges, inclusivity and accessibility, and institutional governance and sustainability. Instead of meta-analysis, narrative thematic synthesis was chosen because the included studies had varied study design, outcome measures, and reporting formats.

In line with the instructions of qualitative synthesis, there were no formal inter-coder reliability statistics (e.g., Cohens kappa). Nevertheless, the application of independent parallel coding as well as systematic consensus discussions is generally accepted in systematic reviews and guarantees the reliability and consistency of the analysis.

The conflicting results of the studies were not ruled out but analyzed in terms of the quality of the methods, the context of the studies, and the characteristics of the sample. This method allowed discovering convergent and divergent patterns. Analytical procedures were strongly related to the conceptual framework of the study, and thematic synthesis was both informed by the empirical patterns and theoretical constructs. The entire coding framework and the examples are detailed in Supplementary File D.

3.2.5. Reporting the Findings

In the final step of the review, the extracted data were organized according to the main themes that emerged from the analysis. These themes were clearly linked to the review questions and the theoretical framework to ensure consistency and coherence. Each theme presented its findings with direct reference to the studies from which the evidence originated, allowing readers to trace the results back to their original sources.

Descriptive statistics, summary tables, and conceptual diagrams were used to illustrate the characteristics of the included studies and to demonstrate the relationships between key opportunities,

challenges, and governance issues. The use of descriptive statistics was limited to summarizing study characteristics and general patterns in the literature; it was not used for inference or statistical testing.

In order to make sure that the theoretical framework would be a guide of analysis and not a justificatory role, we used theory-informed codes based on sociocultural and self-regulated learning views. The coding of studies was based on the features of scaffolding (e.g., its presence, compatibility with the learner ZPD, fading as the learner becomes more competent) and self-regulation indicators (e.g., goal-setting, monitoring, strategic use vs. dependency, evaluation of AI outputs). These codes based on the underpinning theories were used in conjunction with inductive codes. These were used to interpret the themes during synthesis to clarify why benefits and challenges arose and what could be identified as the discrepancies between tool use and theoretical principles.

The implications of these findings for sustainable and inclusive higher education are discussed in the following sections. Furthermore, key research gaps were identified, and practical, evidence-based recommendations were offered for educators, institutions, and policymakers.

4. Results

4.1 Study Selection

The selection of the studies was based on the PRISMA 2020 principles and is presented in Figure 3. The first search of the database occurred in Scopus, Web of Science, and Embase and resulted in $n = 650$ records. The elimination of duplicates resulted in $n = 250$.

After initial screening on the basis of the publication years, the type of document and its topic, $n = 210$ records were filtered out. The other $n = 190$ papers were then filtered through titles and abstracts leading to the removal of $n = 145$ records that were unrelated to AI-based writing aids in the higher education context and did not serve the purpose of the review.

$N = 45$ full-text articles underwent the evaluation of eligibility. Among these, $n = 17$ articles were excluded because of the inability to access the full texts and $n = 1$ study excluded because of the population criteria (i.e. not higher education). In the end, a total of 27 studies were included in the final analysis, as they fulfilled all inclusion criteria.

Table 1 presents a comprehensive analysis of the causes of exclusion at each phase and allows complete openness and repeatability of the selection of the study sample.

Table 1: Reasons for Study Exclusion

Stage	Reason for Exclusion	Number
Initial filtering	Year, document type, subject mismatch	210
Title/Abstract	Not AIW-related / not higher education	145
Full-text	Full text unavailable	17
Full-text	Wrong population	1

The literature search plan was formulated based on a combination of key words and Boolean operators, in order to get the studies associated with artificial intelligence writing aids in higher education. The keywords were the following: AI writing tools, automated writing help and the names of specific tools (e.g., ChatGPT), along with keywords connected to education (e.g., higher education, university, academic writing). In Scopus, Web of Science and Embase, searches were done using consistent Boolean logic (AND/OR), truncation and field restrictions where necessary.

The complete PRISMA 2020 flow diagram illustrating the study selection process is presented in **Figure 3**,

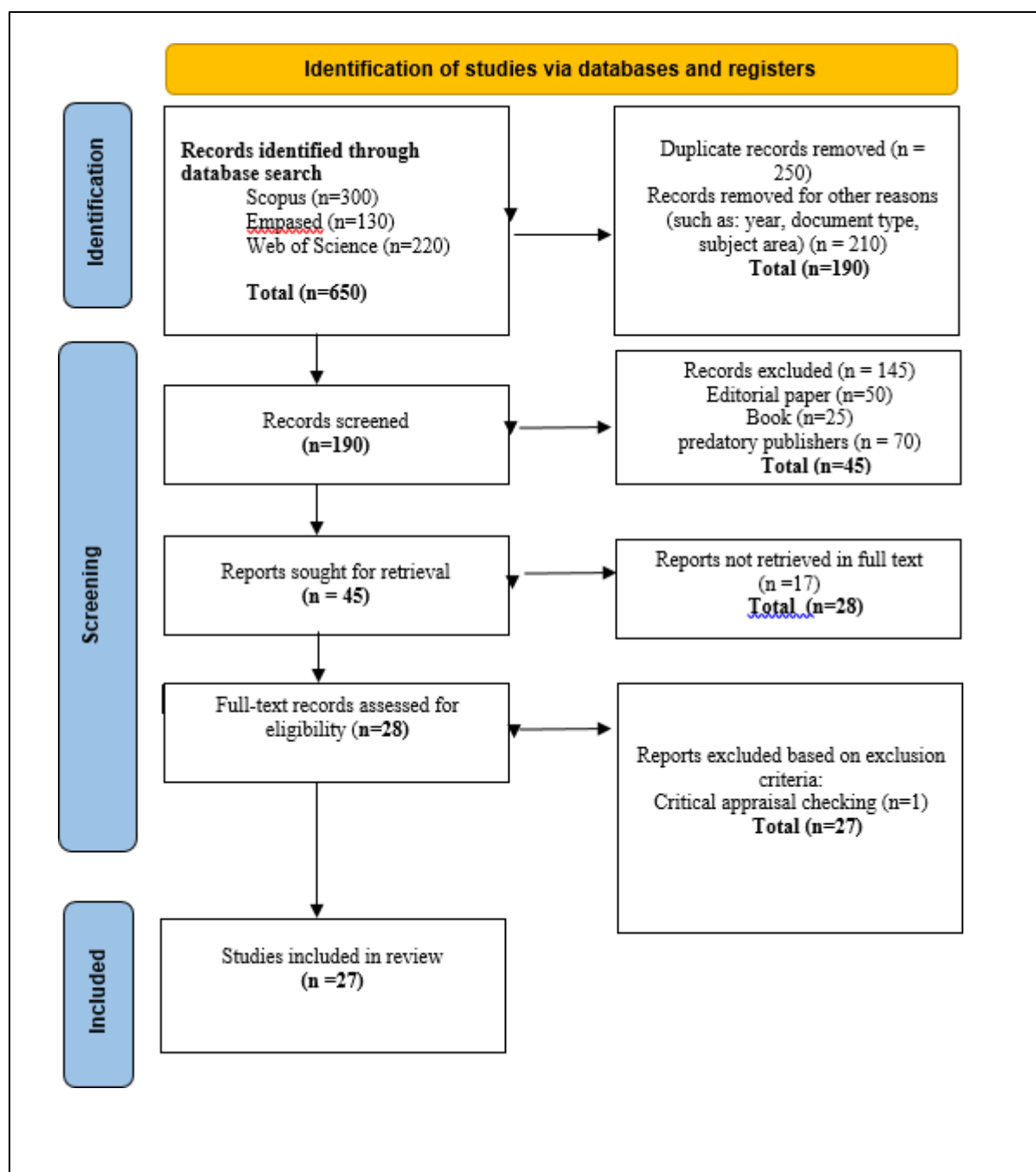


Figure 3. PRISMA Guideline

To enhance the transparency and enable the tracing of the synthesis, an evidence matrix that summarizes the nature of all the included studies is presented in Supplementary File C. The matrix specifies the setting of each study, research design, type of AIW tool used, contributions to the topic area, and quality appraisal scores, allowing the reader to trace the findings of each study to the synthesis.

An evidence matrix was created to show a systematic compilation of extracted data within studies categorized by study characteristics, type of tools, thematic coding, and quality of assessments (see Supplementary File D). Feedback that AIW tools gave in real-time was found to correlate with better quality of writing and revision. Findings from multiple past studies provide strong evidence to support this claim. Issues of academic honesty, such as fears of plagiarism and authorship, were also commonly reported in the literature (strong evidence). Conversely, the effect of AIW tools on critical thinking was inconsistent, and some studies depicted less cognitive engagement, and other studies indicated some

potential benefits when applied in a reflective manner (moderate evidence). Problems of accessibility and inclusivity were defined in a fewer number of studies, especially concerning the cost and infrastructure obstacles (limited evidence).

Table 2. Key Finding for Each Theme

Theme	Key Finding	Evidence Strength
Pedagogical Benefits	Improved writing quality	Strong
Ethical Challenges	Academic integrity concerns	Strong
Inclusivity	Access inequality	Limited

The classification of evidence strength is informed by the number of supporting studies and their methodological quality, as detailed in Supplementary File D.

4.2. Characteristics of Included Studies

4.2.1. Publication Year and Study Design

A total number of 27 studies were included. These studies represent a broad range of geographical regions, research methods, and disciplinary contexts. The studies used diverse research designs, including systematic and mapping reviews, survey-based studies, mixed-methods research, classroom-based interventions, as well as qualitative and quantitative approaches. Table 1 shows a summary of the studies, to illustrate the main types of evidence in the literature (e.g., Imran & Almusharraf, 2023; AlAfnan et al., 2023; Gustilo et al., 2024; Costa et al., 2024; Zhao et al., 2024; Kővári, 2025).

1

Table 3. Summary of Studies

Author(s) & Year	Study Type / Method	Sample / Context	AI Tools / Technologies Examined	Key Findings (Opportunities, Challenges & Sustainability)
Almashy et al. (2024)	Experimental comparative study (pre/post-test)	ESL/EFL learners divided into four groups (ChatGPT, Grammarly, Google Translate, control)	ChatGPT; Grammarly; Google Translate (CALL tools)	All AI tools significantly reduced writing errors, but tool-specific strengths were observed: ChatGPT improved verb forms and subject-verb agreement; Grammarly excelled in prepositions; Google Translate improved articles and capitalization. Findings support targeted, pedagogically guided AI use rather than general reliance, highlighting AI as a supportive tool for sustainable language learning.
Aqeel et al. (2025)	Narrative review / policy-oriented analysis	Higher education in Pakistan; applied and social sciences	ChatGPT; Grammarly; QuillBot; Turnitin	AI writing tools enhance grammar, content generation, paraphrasing, and research support, especially for under-resourced and non-native English-speaking students. Challenges include plagiarism, over-reliance, misinformation, bias, and weak policy frameworks. The study calls for ethical AI governance, institutional policies, AI literacy, and balanced human-AI collaboration to sustain academic integrity.
Aulia & Yamin (2024)	Systematic review (PRISMA 2020)	15 studies from multiple countries; EFL secondary & higher education	ChatGPT (writing assistance)	Identifies three major limitations of ChatGPT in EFL writing: (1) inability to support the entire writing process, (2) over-reliance and threats to originality, and (3) inaccurate or misleading feedback. While useful for idea generation and language support, ChatGPT requires strong teacher guidance and ethical oversight to ensure sustainable academic writing practices.
Bates, Cobo, Mariño & Wheeler (2020)	Editorial / conceptual synthesis (special issue introduction)	Higher education systems internationally	Artificial Intelligence (general educational AI applications)	Despite strong expectations, there is limited empirical evidence that AI has significantly improved teaching and learning outcomes in higher education. Most demonstrated impact lies in learning analytics and prediction, rather than pedagogy or writing quality. The paper cautions against technological hype, arguing for human-centred, augmentation-based AI use rather than replacement. Provides a critical foundation for understanding the gap between AI promises and sustainable academic realities.
Bin-Nashwan, Sadallah & Bouteraa (2023)	Quantitative empirical study (survey + SEM)	702 academics and researchers worldwide (ResearchGate & Academia.edu users)	ChatGPT	ChatGPT use is positively influenced by time-saving benefits, academic self-efficacy, self-esteem, perceived stress, and e-word of mouth. Academic integrity negatively affects adoption, yet also moderates usage behavior. Findings highlight the tension between

				productivity gains and integrity risks, stressing the need for clear ethical frameworks and institutional governance.
Bui & Tong (2025)	Mixed-methods empirical study (survey + semi-structured interviews; thematic analysis)	114 English-major seniors at a Vietnamese university	ChatGPT; Grammarly; QuillBot; AI writing tools	Students perceive AI writing tools as useful for grammar correction, idea generation, and writing efficiency, especially for non-native English speakers. However, major concerns include academic dishonesty, over-reliance, blurred authorship, and difficulty detecting AI-generated text. The study stresses the need for clear institutional guidelines, ethical training, and responsible AI use to preserve academic integrity.
Cheuquepán-Maldonado & Chourio-Acevedo (2025)	Quasi-experimental study; text-mining analysis; Mann–Whitney U tests	160 Spanish-speaking vocational education students in an online communication literacy course (Chile)	Generative AI embedded in an LMS discussion forum (AI-powered feedback system)	Use of generative AI–supported forums did not result in statistically significant improvements in textual readability compared to traditional forums (FHR $p = 0.233$; SPP $p = 0.262$). While AI assisted with idea generation and revision, it did not measurably enhance surface-level writing quality. The authors suggest that AIW benefits may lie beyond readability—such as coherence, cohesion, and higher-order writing skills—and emphasize the need for multi-dimensional assessment frameworks. Sustainable AIW integration requires pedagogically guided use and broader evaluation metrics, not reliance on superficial writing indicators alone.
Chalkiadakis et al. (2024)	Systematic review	Global higher education context; focus on students with disabilities and inclusive education	Artificial Intelligence (AI); Virtual Reality (VR); adaptive learning systems	AI-driven technologies support personalised learning, accessibility, and inclusion, particularly for students with disabilities. Challenges include high implementation costs, technical barriers, algorithmic bias, privacy concerns, and limited teacher readiness. Although not focused exclusively on writing tools, the review provides important sustainability and equity perspectives relevant to AI-assisted academic practices, highlighting the need for ethical governance and inclusive design.
Costa, Ntsobi & Mfolo (2024)	Systematic mapping review (research article; review-based)	Literature from 22 countries; higher education academic writing context (with focus on South Africa)	ChatGPT; Generative AI (GAI) writing tools	ChatGPT significantly enhances productivity, language support for non-native English speakers, and immediate feedback in academic writing. However, it presents serious challenges related to plagiarism, fabricated references, accuracy, over-reliance, and erosion of critical thinking. The study proposes the OTHA Framework (Openness, Transparency, Honesty, Accountability) to guide ethical and sustainable AI use. Recommends institutional

				policies, disclosure of AI use, ethical training, and monitoring to ensure responsible academic writing practices.
Chanpradit et al. (2024)	Qualitative descriptive case study; text analysis using Keck's (2014) and Nabhan et al.'s (2021) frameworks	30 Scopus-indexed research abstracts paraphrased by AI tools	QuillBot; Paraphrasing Tool (standard modes)	Both AI paraphrasing tools primarily rely on synonym substitution and sentence restructuring. QuillBot tends toward minimal to moderate revision, while Paraphrasing Tool shows more moderate revision, with limited instances of substantial rewriting. Neither tool consistently achieves deep conceptual transformation, indicating surface-level paraphrasing. The study warns that AI paraphrasing tools may reduce genuine paraphrasing skill development and pose academic integrity risks if overused, recommending teacher-guided use, AI as a secondary aid, and explicit instruction in ethical paraphrasing for sustainable academic writing.
Deep & Chen (2025)	Narrative review (SANRA-guided)	20 peer-reviewed studies (2023–2025); higher education students and faculty	ChatGPT; Grammarly; QuillBot; Zotero; Mendeley	AI writing tools enhance writing fluency, grammar, feedback speed, self-directed learning, and cognitive load reduction, but risk over-reliance, weakened critical thinking, plagiarism, loss of originality, and integrity violations. Effective integration depends on pedagogical scaffolding, ethical guidance, transparency, and curriculum redesign. The review emphasizes balanced human–AI collaboration for sustainable academic writing development.
Dergaa, Chamari, Zmijewski & Ben Saad (2023)	Quasi-qualitative literature review	Academic research and publishing context; higher education and scientific writing	ChatGPT (GPT-3), NLP-based AI tools (BERT, RoBERTa, QuillBot, Paperpal)	AI-based writing tools can dramatically improve academic efficiency, including summarization, idea generation, literature scanning, and drafting support. However, they pose significant threats to authenticity, credibility, authorship transparency, plagiarism, and misinformation, including the risk of AI-generated fake science. The study strongly argues that AI cannot replace human researchers, must never be credited as an author, and requires strict transparency, human verification, ethical governance, and accountability. Sustainable academic practice depends on human-AI collaboration, not substitution.
Duah & McGivern (2024)	Qualitative empirical study (semi-structured interviews; inductive thematic analysis)	4 university students and 1 lecturer; UK higher education	ChatGPT; Generative AI tools	Generative AI blurs authorial identity and traditional definitions of academic misconduct, especially in summative assessment. Students are generally more open to AI as a support tool, while educators are more cautious. Both groups highlight policy gaps and ethical uncertainty. The study stresses the urgent need for clear institutional AI usage policies and dialogue between students and faculty to sustain academic norms.

Ely & Rezvani Rad (2024)	Case study / classroom-based empirical study	Undergraduate engineering students; technical writing course (USA)	ChatGPT; Grammarly; Hemingway (generative & non-generative AI)	AI tools enhance grammar, structure, feedback speed, and revision quality, supporting professional writing development. However, AI-generated text often lacks voice, nuance, contextual accuracy, and numerical reliability, and may constrain creativity. The study recommends structured AI literacy modules, clear usage boundaries, and ethical instruction for sustainable integration in writing-intensive courses.
Gallent-Torres, Zapata-González & Ortego-Hernando (2023)	Narrative literature review / conceptual analysis	Higher education institutions globally; students, faculty, and institutions	ChatGPT; Humata.ai; Sudowrite; Generative AI tools	Generative AI offers opportunities such as instant feedback, adaptive learning, accessibility, and resource generation, but raises major ethical challenges related to reliability of information, transparency, data privacy, sustainability, and new forms of plagiarism and authorship impersonation. The authors argue that sustainable AIW integration requires clear ethical guidelines, institutional integrity policies, and pedagogical practices that guide responsible use.
Gustilo, Ong & Lapinid (2024)	Cross-sectional empirical survey grounded in TAM	100 graduate students and faculty across 10 disciplines; Philippines	ChatGPT; Grammarly; Google Translate; QuillBot (Algorithmically-Driven Writing Tools)	Educators perceive high usefulness of AI writing tools for teaching, learning, and assessment, but face barriers including limited access, low AI self-efficacy, and serious concerns about academic integrity, creativity, and output quality. The absence of explicit institutional policies intensifies ethical ambiguity. The study calls for collaborative governance, AI literacy, and policies that privilege human intelligence over automation to ensure sustainable academic practice.
Imran & Almusharraf (2023)	Systematic literature review	30 peer-reviewed studies; global higher education	ChatGPT (GPT-3.5 & GPT-4)	ChatGPT improves writing efficiency, idea generation, language accuracy, and collaboration. However, significant challenges include plagiarism, fabricated citations, assessment integrity, and lack of clear institutional policies. The authors emphasize AI as a supportive aid rather than a replacement for human writing, calling for policy reform and AI literacy training to ensure ethical and sustainable adoption.
Kadwa (2025)	Conceptual / analytical research article (ethical and pedagogical analysis)	Higher education institutions; written assignments; multilingual and second-language contexts	Generative AI tools (e.g., ChatGPT)	AI-based writing tools provide valuable linguistic and structural support, particularly for second-language learners, but raise serious concerns related to academic integrity, transparency, bias, misinformation, equity, authorship, and assessment validity. Traditional written assignments are especially vulnerable. The study recommends process-oriented assessments, reflective portfolios, oral

				defenses, mandatory AI disclosure, and AI literacy education to ensure ethical, fair, and sustainable academic practices.
Koos & Wachsmann (2023)	Conceptual / analytical study (legal education focus)	Higher education, particularly law students and academic examinations	ChatGPT; GPT-4	AI tools support idea generation, structuring, language quality, and productivity, benefiting students with language barriers. However, risks include plagiarism, hallucinated sources, reduced creativity, and erosion of critical thinking. The paper emphasizes ethical awareness, regulation, assessment redesign, and skill development to ensure responsible and sustainable AI-assisted academic writing.
Kotsis (2025)	Qualitative thematic synthesis (review) of 33 peer-reviewed studies (2022–2025)	Higher education academic writing contexts; interdisciplinary	ChatGPT (LLM-based generative AI)	ChatGPT can function as a writing coach, supporting brainstorming, drafting, revision, and multilingual learners. However, risks include algorithmic dependence, fabricated content, plagiarism, unequal access, and erosion of authorship. The study emphasizes AI literacy, transparency, disclosure, equitable institutional policies, and critical pedagogy as prerequisites for responsible and sustainable AI-assisted academic writing.
Kovari (2025)	Opinion / conceptual paper (best-practice framework)	Higher education institutions; educators and students	ChatGPT; generative AI tools	Highlights serious risks of AI-induced plagiarism, hallucinated references, lack of originality, and weakened critical thinking. While ChatGPT enhances accessibility, engagement, and personalized feedback, its misuse threatens academic integrity. The paper proposes best practices such as transparent AI disclosure, adaptive and reflective assessments, oral exams, creative tasks, peer review, and strengthened plagiarism detection. Emphasizes ethical governance, AI literacy, and human oversight as essential for sustainable academic writing.
Mokhine-Martins, Deka & Costa (2024)	Systematic mapping review protocol (PRISMA-guided)	African communities; culturally grounded well-being research	<i>Not AI-writing focused</i> (conceptual relevance only)	Although not directly focused on AI-based writing, this protocol contributes a sustainability and ethics lens by integrating Ubuntu philosophy with positive psychology, emphasizing community, equity, inclusion, and cultural relevance. Its principles support human-centred, ethically grounded approaches applicable to AI governance in higher education, particularly in Global South contexts.
Oates & Johnson (2025)	Mixed-methods empirical study (quasi-experimental; rubric-based assessment; focus	10 Master's students in biomedical science programmes; Leeds Beckett University (UK)	ChatGPT	ChatGPT demonstrated strong grammatical accuracy and structural coherence, but did not improve overall essay marks. However, when students critically evaluated AI-generated essays, their critical evaluation scores improved significantly. Students preferred writing their own essays and viewed ChatGPT as a scaffolding and

	group; questionnaire)			evaluative tool rather than a writing substitute. Sustainable AIW use lies in pedagogical designs that promote critical thinking, human oversight, and reflective engagement.
Sarin & Kimkong (2024)	Conceptual / policy analysis	Cambodian higher education context	ChatGPT; Grammarly; QuillBot; Google Translate; LLMs	Highlights urgent need to develop or revise academic integrity policies due to rising risks of plagiarism, cheating, reduced critical thinking, privacy issues, and unreliable AI outputs. Argues that banning AI is unsustainable; instead, institutions should adopt policy-driven, ethical, and pedagogically guided AI use to support long-term academic sustainability, particularly in developing contexts.
Sterling, Ye, Ying & Chen (2025)	Qualitative exploratory study / reflective exercises	International undergraduate and postgraduate students; computing disciplines	ChatGPT (GPT-4); Claude 3.5; Llama 3 (LLMs)	Generative AI helps international students overcome language barriers, blank-page syndrome, and drafting difficulties, improving fluency and efficiency. However, overuse leads to homogenised writing, loss of authorial voice, exaggerated claims, and authenticity concerns. Sustainable AIW use requires careful prompt design, reflective practice, transparency, and explicit instruction on preserving personal voice, positioning AI as a scaffold rather than a substitute.
Turingan (2025)	Sequential explanatory mixed-methods study (survey + interviews)	326 tertiary students (quantitative) + 12 interviewees; higher education	ChatGPT; Grammarly; QuillBot; AI writing tools	Students generally perceive AI writing tools as useful for improving efficiency, organisation, and grammar, supporting perceived usefulness and ease of use. However, major concerns include over-reliance, plagiarism, ethical misuse, and inconsistent institutional AI policies. Significant differences in attitudes were found based on sex, academic year level, and frequency of use. The study emphasizes the need for AI literacy programs, ethical frameworks, and institutional guidelines to ensure sustainable and responsible AI-based academic writing.
Zhao, Cox & Chen (2025)	Empirical quantitative study (survey informed by AI-literacy framework)	124 university students with disabilities; UK higher education	ChatGPT; generative AI chatbots; rewriting tools	Generative AI is widely used by students with disabilities for proofreading, rewriting, summarising, and academic writing support. Benefits include accessibility and learning support, while major concerns include inaccuracy, hallucinations, academic integrity risks, subscription costs, and lack of institutional guidance. Students expressed strong demand for AI literacy training, inclusive policies, and involvement in AI policymaking, highlighting AI's potential for inclusive and sustainable education when governed ethically.

4.2.2 Methodological Distribution of Included Studies

Across the final dataset of 27 included studies, systematic and narrative review papers account for approximately 30% of the literature ($n = 8$). These review studies were intentionally included to capture synthesised evidence, theoretical discussions, and ethical debates in a rapidly developing field. Given the early stage of research on AI-based writing tools in higher education, review papers play an important role in identifying overall trends, conceptual gaps, and emerging sustainability concerns. Empirical studies were mainly used to support conclusions related to learning outcomes and practical use.

Among the empirical studies, cross-sectional survey designs represent 22% of the dataset ($n = 6$), indicating a strong focus on perceptions, attitudes, and adoption behaviours of AI-based writing tools. Qualitative studies account for 19% ($n = 5$) and provide deeper insights into user experiences, ethical concerns, and pedagogical implications. Mixed-methods studies also make up 19% ($n = 5$), reflecting increasing efforts to combine quantitative and qualitative evidence. In contrast, experimental or quasi-experimental studies are relatively limited, representing only 13% of the included studies ($n = 3$).

Overall, this distribution suggests that research on AI-based writing tools in higher education is still largely exploratory and descriptive in nature. There remains a clear lack of rigorous intervention-based studies that examine causal effects and long-term learning outcomes. As a result, claims about pedagogical effectiveness are interpreted with caution when they are based mainly on cross-sectional or self-report data. Figure 3 illustrates the distribution of study designs across the included literature.

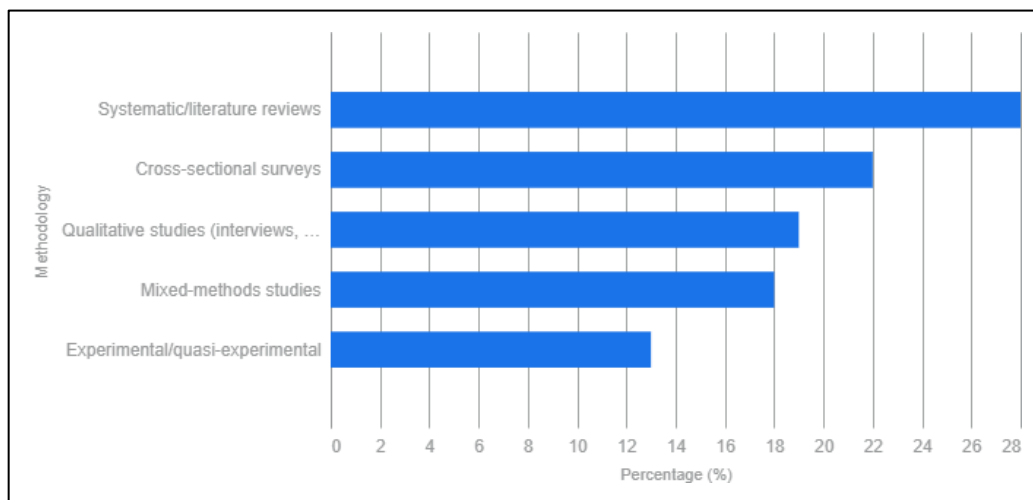


Figure 3: Methodology Distribution

4.2.3. Geographical Distribution

The included studies show an uneven geographical distribution across global regions. Around 32% of the studies ($n = 9$) were conducted in North America, followed by Europe with 28% ($n = 8$) and the Asia–Pacific region with 24% ($n = 6$). Research from the Middle East accounts for 11% of the studies ($n = 3$), while Africa is represented by only one study (3%). In addition, a small number of studies were conducted across multiple countries or did not specify a single geographical context, most often in large-scale review or conceptual papers. Overall, this pattern shows that research on AI-based writing tools in higher education is largely concentrated in the Global North. In contrast, evidence from under-resourced and Global South regions remains limited. This imbalance is important, as several review and multi-country studies included in this review highlight the need for broader geographical coverage to better reflect differences in educational systems, infrastructure, language contexts, and equity conditions. The findings therefore point to a clear geographical research gap and underline the need for more empirical studies conducted in diverse cultural, educational, and socio-economic settings to support more

inclusive and sustainable academic practices. Percentages are rounded to the nearest whole number (see Figure 4).

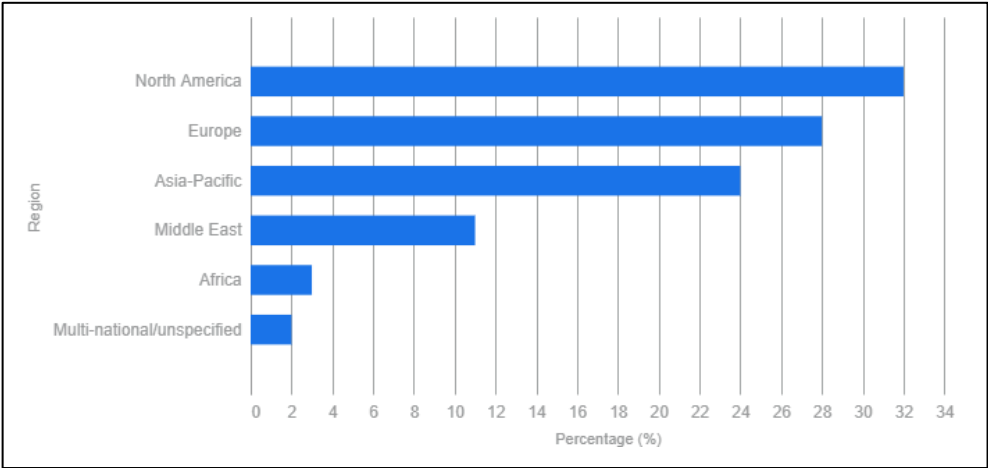


Figure 4, Geographical Distribution

4.2.4 AI-Based Writing AI Tools Examined

Regarding the technologies examined, ChatGPT and related large language models (e.g., GPT-3.5 and GPT-4) dominate the literature, appearing in approximately 67% (n = 18) of the included studies. In addition, around 30% (n = 8) of studies refer more broadly to multiple or unspecified generative AI tools, reflecting a growing focus on general AI-assisted writing practices rather than single platforms.

By comparison, non-generative or task-specific writing tools are discussed less frequently, including Grammarly (18%, n = 5), QuillBot (12%, n = 3), and other specialist academic writing assistants such as Writefull, Scite, and Jasper (8%, n = 2). As several studies examined more than one tool, these categories are not mutually exclusive (see Figure 5).

Overall, this distribution shows that while ChatGPT has received considerable research focus, other AI-based writing tools are relatively less researched, which presents an opportunity for more diversified research in future studies. See figure 5.

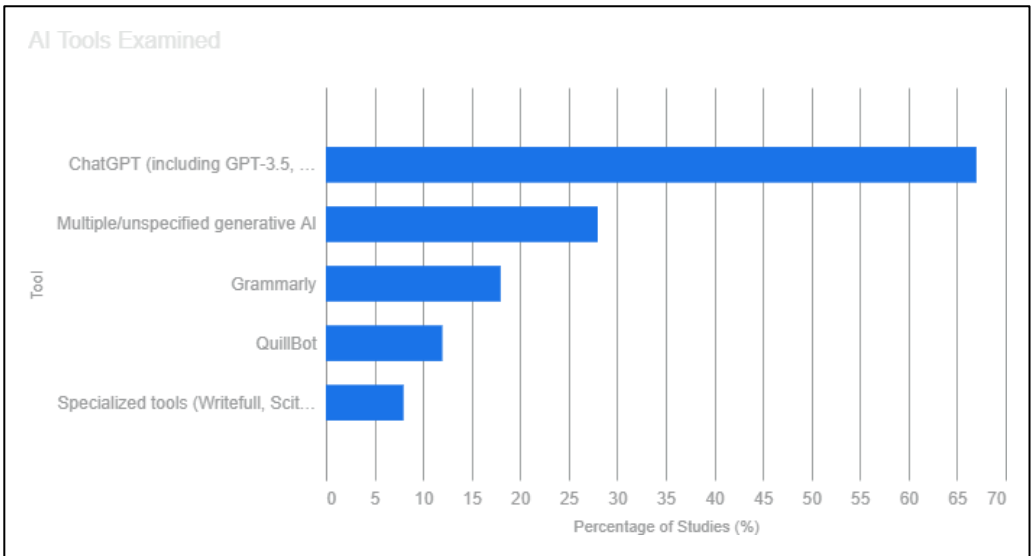


Figure 5. General AI Writing Tools

4.3 Methodological Quality Overview

The methodological quality of the included studies was assessed using the Mixed Methods Appraisal Tool (MMAT, 2018). Overall, the quality of the evidence ranged from moderate to high, with a smaller number of studies assessed as low quality due to limitations in sampling, study design, or analytical transparency.

Most studies met the MMAT criteria related to clarity of research questions and alignment between research aims, methods, and outcomes. Studies rated as high quality typically provided clear methodological descriptions, appropriate data collection procedures, and well-justified analytical approaches. Moderate-quality studies often relied on self-report data or cross-sectional designs, which limited causal inference but still offered valuable descriptive and contextual insights. A small number of low-quality studies showed weaknesses such as limited sample justification or insufficient detail in data analysis.

Methodological quality was not used as an exclusion criterion. Instead, MMAT assessments were used to inform the level of confidence placed in different findings during synthesis. Evidence from higher-quality studies was given greater weight when drawing conclusions, particularly for claims related to pedagogical effectiveness and learning outcomes, while findings from lower-quality studies were interpreted more cautiously.

All the scores of the included studies in terms of the detailed evaluation of MMAT (2018) quality are provided in Supplementary File C, which enables readers to assess the methodological quality of the study on their own. During the main analysis, the quality of the studies was taken into account when interpreting the results whereby more consideration was given to patterns that were supported by higher-quality studies. Poor quality studies were not discarded but were viewed with caution especially where the results were not backed by the high-quality studies. Quality appraisal using MMAT was conducted independently by two reviewers, with discrepancies resolved through discussion and consensus.

4.4. Thematic Synthesis of Findings

Across the 27 included studies, pedagogical benefits were reported in 18 studies, most commonly in relation to writing efficiency, feedback, and language support. Ethical and integrity challenges were identified in 21 studies, making this the most consistently reported theme, particularly in relation to plagiarism, over-reliance, and authorship ambiguity. Institutional governance and sustainability concerns were raised in 16 studies, with emphasis on policy development, AI literacy, and assessment redesign. Inclusivity and accessibility were explicitly addressed in 6 studies, primarily focusing on support for non-native English speakers and students with disabilities.

Together, these patterns indicate that current knowledge about AIW tools in higher education is shaped largely by early-adopting institutions in the Global North, relies heavily on self-report and cross-sectional designs, and focuses predominantly on generative systems such as ChatGPT.

Based on the thematic synthesis, four higher-order themes were identified: pedagogical benefits, ethical and integrity challenges, inclusivity and accessibility, and institutional governance and sustainability (Figure 6). Details of the coding framework and theme extraction process are provided in Supplementary File D, supporting transparency of the synthesis. Initial open coding generated descriptive sub-themes (e.g., writing efficiency, academic integrity, accessibility, governance), which were subsequently grouped into higher-order themes through iterative comparison and refinement. This hierarchical structuring ensured conceptual coherence between the extracted data and the final thematic framework.

From a theoretical perspective, these patterns can be interpreted through sociocultural and self-regulated learning frameworks. AIW tools function as scaffolding mechanisms when used strategically, enabling learners to refine their work within their developmental capacity. However, when these tools replace

rather than support cognitive processes, they may undermine learning. Across themes, benefits were most evident where tools operated as scaffolds with gradual fading and where learners demonstrated strategic self-regulation, whereas risks were more pronounced in contexts characterized by automation without fading and dependent use.

Conflicting findings across studies were retained and interpreted in relation to methodological quality, study context, and sample characteristics. This approach enabled the identification of both convergent and divergent patterns within the evidence base.

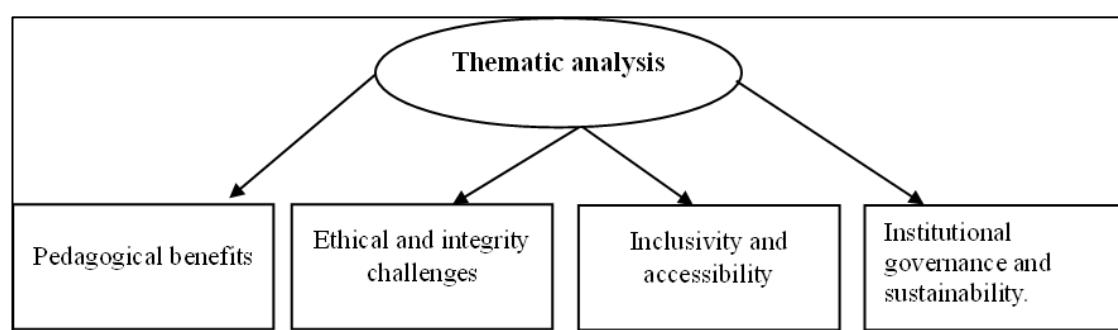


Figure 6. Higher-Order Themes Derived from Thematic Synthesis of the Included Studies

4.4.1. Theme 1: Pedagogical Benefits and Learning Outcomes

The literature documents several pedagogical advantages of AIW tools, with important caveats on instructional design and implementation.

1. Immediate Feedback and Writing Improvement

Perhaps the most highly cited advantage of using AIW tools is their capability to provide instant feedback and correction on grammar, coherence, structure, and style. Many studies show that students, especially multilingual learners, use AIW tools to refine writing drafts and improve accuracy and coherence. For example, Angelos et al. (2024) found that 89% of the Saudi EFL learners in their study reported improved academic writing skills after integrating the use of ChatGPT into the process of writing revision. They also found enhanced confidence among the learners in expressing complex ideas.

2. Generation of Ideas and Creativity Enhancement

AIW tools have been found to be useful in the brainstorming stage. For example, AlAfnan's (2023) study found that ChatGPT can generate multiple examples, alternative phrasings, and useful conceptual frameworks to help students and instructors as a starting point for their written work. Konstantinos (2025) synthesized 33 studies and found that ChatGPT was useful as a writing coach to help students overcome writing anxiety and explore different approaches for academic writing. This scaffolding, in the form of content ideas, appears especially valuable for novice writers and those unfamiliar with academic writing in their disciplines. However, researchers caution against overreliance on AI-generated ideas as they may constrain rather than expand creative thinking if students adopt suggestions uncritically rather than using them as spring boards for original thought (Ely & Rad, 2024).

3. Scaffolding and Personalized Learning Support

Several studies position AIW tools as cognitive scaffolds that can be gradually withdrawn as learners develop competence. King et al.'s (2024) mapping review across 22 countries found consistent evidence that generative AI provides on-demand tutoring, offering explanations of writing conventions, modeling revision strategies, and responding to student queries, functions traditionally limited by instructor availability and class size. Kotari (2025, p.7) suggests that "another pedagogical method that AI has

changed in collaborative environments involves project-based learning. It requires learners to work in groups and solve complex real-world problems with continuous feedback, recommendations, and resources provided by these AI systems.” This illustrates the learning support provided by AI systems to student learning, including for writing tasks.

4. Efficiency and Workflow

Past studies consistently found that AIW tools helped both lecturers and students save time when carrying out writing tasks such as drafting outlines, organizing citations, and summarizing materials. This helps free up cognitive resources for higher-order thinking skills such as analyzing, evaluating, and synthesizing past studies, making inferences, and drawing conclusions. For example, Gustilo et al.'s (2024) survey of 100 Filipino lecturers and graduate students found that perceived usefulness was the strongest predictor of adoption intention, with efficiency benefits outweighing concerns.

However, critics have argued that efficiency is achieved at the expense of learning. Critics within the reviewed literature argue that the productive challenges inherent in writing, which are part of the process for developing writing skills such as dealing with ambiguity, revising repeatedly, and developing voice, may be shortchanged when AIW tools make writing "too easy" (Duah & McGivern, 2024). This tension underscores the need for pedagogical intervention. Efficiency is pedagogically valuable when it eliminates busy work, but it is problematic when it bypasses learning processes.

5. Development of Critical Evaluation Skills

Recent studies suggest that integrating AIW tools into pedagogical practices could enhance critical thinking rather than diminish it. A classroom-based study by Oates and Johnson (2025) showed that teaching students to evaluate ChatGPT outputs for accuracy, bias, and appropriateness would lead to strong evaluation and argumentation skills. This aligns with calls for AI literacy to be taught to students as part of the aim to promote critical and ethical use of AI tools in tertiary education.

4.4.2. Theme 2: Ethical and Academic Integrity Challenges

Though pedagogical benefits emerge from the positive findings shown by past literature, there is an ongoing debate on ethical and integrity concerns in the use of AIW tools for academic writing.

1. Plagiarism and Paraphrasing Concerns

Just by providing prompts and questions, AIW tools such as ChatGPT can generate essays that are coherent and sophisticated. Students can pitch the generated essays to a level they require and further prompt the AIW tool to refine the writing. The capacity of AIW tools to produce texts for a multitude of topics and purposes has resulted in students taking this shortcut to academic writing.

2. Authorship and Attribution Ambiguities

The issue of authorship is also a primary concern. To what extent can AIW tools be used for a student to still be able to claim authorship of a piece of writing? When ChatGPT generates an outline, refines the arguments, and edits the writing, who is the author? These are concerns for lecturers, students, and higher education governing bodies. There appears to be a contradiction in the understanding of the use of AI among lecturers and students. Duah and McGivern's (2024) qualitative interviews with UK students and lecturers indicate that students described AI as a tool, much like Google, but lecturers tend to view the use of AI as academic dishonesty, as it undermines authenticity and student ownership of writing.

3. AI Detection Challenges

The literature documents growing scepticism about the use of AI-detection tools. These tools are often seen as producing false positive or false negative rates. The different tools also often show discrepancies in the detection of AI-generated text, particularly for non-native English speakers whose writing may be flagged as AI-generated due to stylistic patterns (Kővári, 2025). AlAfnan et al. (2023) found that even straightforward paraphrasing requests to ChatGPT rendered text undetectable by current tools. Thus, the reliability and trustworthiness of these AI-detection tools are highly doubted. However, Kővári's (2025) comprehensive review notes that detection tools may still play transitional roles in institutional responses, particularly when combined with process-based verification methods.

4. Misinformation, and Fabrication

Generative AI has been known to generate fabricated information or what is termed as “hallucinations”. The generated content sounds plausible but actually contains fabricated data, false citations, made-up references, and false claims. A number of studies have shown that students unknowingly incorporate these fabricated references and content into their academic writing (AlSofi, 2024; Kotsis, 2025). When students do not fact-check their AI-generated content, there is the likelihood that they will accept this confident-sounding output generated by AIW tools. This problem underscores the need to teach AI literacy in tandem with teaching writing skills. Table 2 summarizes the challenges in academic integrity and proposed strategies to address these challenges.

Table 2. Academic Integrity Challenges and Proposed Mitigation Strategies

Challenge	Evidence	Proposed Mitigation Strategies	Limitations
Paraphrasing plagiarism	ChatGPT-produced text can bypass similarity checkers (AlAfnan et al., 2023)	• Redesign assessments (contextual, authentic writing tasks) • Require process portfolios and drafts • Enforce disclosure of AI use	Detection tools unreliable; requires major pedagogical shifts
Authorship ambiguity	Students more accepting of AI authorship; staff uncertain without policy (Duah & McGivern, 2024)	• Establish institutional authorship policies • Develop transparent AI acknowledgment norms • Integrate academic integrity and AI ethics training	Persistent philosophical disagreement; policies may not align across disciplines
Detection failures	False positives and negatives common in AI detection systems (Kővári, 2025)	• Reduce reliance on AI detection tools alone • Emphasize iterative writing and process assessment • Use oral defenses or viva checks where relevant	Process-based approaches require staff training and time; institutional capacity varies
Data privacy	Risks of student data being exposed to third-party AI platforms (Costa et al., 2024)	• Use institutionally licensed AI tools • Implement robust data governance and privacy protocols • Train students to critically evaluate platform privacy policies	Institutional solutions costly; limited control over commercial AI data practices
Bias and misinformation	AI may generate biased content and fabricated citations (AlSofi, 2024; Kotsis, 2025)	• Teach critical AI literacy and verification skills • Require citation checks and fact validation steps • Encourage diverse, well-designed prompts	Lecturer and student AI skills vary; risk evolves as models update

Note: Evidence derived from included studies; table summarises reported strategies rather than evaluated effectiveness.

4.4.3. Theme 3: Inclusivity and Accessibility

AIW tools present complex implications for educational equity. Though they offer support for multilingual learners, they raise issues of lack of access for marginalized learners, which further exacerbates concerns of inequality in education.

1. Support for Multilingual and Non-Native English Writers

Support for non-native and multilingual speakers has been identified as a strength afforded by AIW tools. This support includes providing corrections in grammar use, using appropriate idioms, lexical items, and academic register (Costa et al., 2024). AISofi's (2024) study of Saudi EFL undergraduates revealed that ChatGPT use positively influenced writing confidence and willingness to attempt complex sentence structures. This would suggest that AI support may reduce anxiety and encourage risk-taking among language learners. Kotsis's (2025) synthesis similarly found AI as especially beneficial for multilingual and novice writers during brainstorming and revision phases.

2. Assistance for Students with Disabilities

AIW tools have also been found to facilitate students with disabilities in their writing journey. A survey by Zhao et al. (2024) found that 124 students with ADHD, dyslexia, dyspraxia, and autism spectrum conditions in UK higher education actively used generative chatbots and rewriting apps to overcome writing-related barriers, organize thoughts, maintain focus and produce coherent drafts. This would point to the assistance that AIW tools provide for writing tasks among students with certain disabilities.

3. Equity Considerations and Access Barriers

Though AIW tools appear to provide access to writing support for students in higher education, the review of past studies show some threats in equity. Gustilo et al.'s (2024) Philippine study found that access barriers such as limited internet connectivity, lack of institutional subscriptions to AIW tools, and low digital literacy could adversely affect students from low socioeconomic backgrounds. In the same vein, Zhao et al. (2024) reported that premium AIW tool subscriptions (e.g., ChatGPT Plus, Grammarly Premium) created a dividing system where affluent students accessed more powerful features while others made do with limited free versions. Costa et al. (2024) found that these access disparities are consistent with existing educational inequalities, where under-resourced institutions that lack infrastructure can only provide limited access to AIW tools, while well-funded institutions can offer premium licenses and comprehensive training. This pattern risks widening rather than narrowing global educational divides.

4. Representation and Cultural Responsiveness

An emerging concern in the literature revolves around AIW tools' capacity for cultural responsiveness. Given that generative AI systems appear to be more in tune with the English language and western academic texts, it may default into western cultural assumptions and norms, thus marginalizing non-western knowledge systems, writing traditions, and writing conventions (Costa et al., 2024). Thus, students from non-Western backgrounds might feel disconnected from the AIW contexts. When students rely on AIW tools to generate examples, discussions, and arguments, they may receive generated content with a dominance of Western culture and perspectives, thus sidelining non-Western viewpoints. This is especially salient in fields such as humanities, arts and social sciences, where diverse and contradicting perspectives are common.

4.4.5 Theme 4: Institutional Governance and Sustainability

The sustainability of AIW tools also depends on lecturer preparedness, policy development and enforcement, and institutional capacity.

1. Policy Development and Institutional Responses

The reviewed studies highlight concerns among lecturers and students about the lack of clear policies and procedures for the use of AIW tools. Duah and McGivern's (2024, p. 9) study in the UK found that “students may take a more laissez-faire approach to GenAI, viewing it as a complimentary tool and part of the writing process. Academics may view it in a more negative light and a tool which is perhaps more likely to jeopardise authorial integrity.” Thus, the absence of clear policies on the use of generative AI systems results in ethical ambiguity and uncertainties among both lecturers and students. Students were unsure of the permissible use of AI, while lecturers showed inconsistent practices in allowing AI use among students. This lack of policy appears to be widespread among higher education institutions. In a similar vein, Costa et al.'s (2024) review found that fewer than 30% of institutions studied had published comprehensive AI use guidelines by mid-2024.

In general, current guidelines in the use of AIW tools require students to disclose how they used generative AI in their writing process (Kővári, 2025). However, collectively, there appears to be a lack of understanding of disclosure requirements and practices. A number of studies adopt the use of the OTHA framework, which encompasses Openness, Transparency, Honesty, and Accountability. This framework emphasizes transparent and explicit acknowledgement of AI-generated content and its use in students' written work. However, how this framework translates into actual practices need to be further investigated and made explicit for lecturers and students.

2. Lecturer Training and Digital Literacy Needs

Lecturer preparedness in incorporating AIW tools and managing students' use of these tools also poses a challenge for their optimal and ethical use. Gustilo et al.'s (2024) survey found that even though Filipino lecturers recognized the usefulness of AIW tools, they reported low self-efficacy in using the tools within the teaching and learning context. They also reported limited understanding of the AIW tools' capabilities and limitations. This study raises several issues, including misconceptions about how AIW tools function and ill-informed policy development. Thus, there is a call for professional development in AI-literacy and AI policy development. Literature suggests that comprehensive lecturer professional development should be carried out and it should include hands-on experience with AIW tools, pedagogical strategies for integrating AIW tools, ethical guidelines for AI use, and AI literacy (e.g., Kővári, 2025; Turingan, 2025).

3. Assessment Redesign for Sustainable Integrity

Rethinking and reformulating assessment practices are key in ensuring ethical and genuine authorship contributions in writing. Rather than prohibiting the use of AIW tools, a more productive approach would be to implement evidence-based strategies (Turingan, 2025). This would include the use of process-based writing, where the writing and revising of the drafts are monitored. Another strategy would be using authentic and contextualized tasks as assignments requiring personal experience, local knowledge, or specific course content are harder to outsource to generative AI tools (Alafnan et al., 2023). Adding an oral component to the writing task could also ensure academic writing honesty. Pairing written work with oral defenses, presentations, or in-class discussions allows for verification of understanding and reduces the chances for AI misuse (Kővári, 2025). Multimodal and collaborative projects could also help maintain integrity in the written work. Assignments incorporating audio, video, visual, or hands-on components or requiring genuine collaboration are more difficult to complete solely through AI tools (Kővári, 2025). In addition, group projects with peer evaluation add social accountability that discourages free riding through AI. Metacognitive reflection could also be used as it

would require students to reflect on their writing process, including AI tool use. This could foster transparency and critical thinking about technology's role in learning (Turingan, 2025).

4. Infrastructure and Resource Considerations

Sustainable AI integration requires institutional investment in infrastructure such as licenses for AI tools, technical support systems, learning management system integration, and ongoing evaluation mechanisms (Gustilo et al., 2024). The literature suggests that ad hoc, individual adoption, creates inequities and confusion. Thus, there is a need for higher institutions to play active roles in ensuring adequate investment to support the use of AIW tools that would benefit learners. Another aspect that needs attention is the environmental cost we pay as a result of the extensive use of AI systems. Large language models require substantial computational resources and energy, raising questions about the carbon footprint of widespread AIW tool use (Richie, 2025). Only a handful of studies mention environmental sustainability, representing a significant gap given higher education's commitment to climate action and SDG alignment.

5. Alignment with SDG 4: Quality Education Goals

While many studies implicitly connect AIW tools to quality education goals, explicit alignment with UN Sustainable Development Goal 4 remains limited. The reviewed literature suggests several potential connections to SDG 4: Quality Education, through encouraging AI literacy as a skill for employment, as well as promoting equity and inclusivity in education. However, currently these connections remain largely aspirational and vague. The gap between AIW tools' potential to advance SDG 4 and actual implementation that realizes this potential represents a critical area for future research and policy development.

Table 3 summarizes the institutional governance dimensions and implementation challenges highlighted by the reviewed studies. These are the challenges related to SDG 4 and the use of AIW tools.

Table 3. Institutional Governance Dimensions and Implementation Challenges

Governance Dimension	Current State	Challenges	Recommendations
Policy Development	Many institutions lack clear AI use guidelines (Duah & McGivern, 2024)	Rapid AI evolution; disciplinary norms differ; philosophical debates on authorship and integrity	Adopt OTHA governance framework; involve students & lecturers in policy design; review and update policies iteratively.
Lecturer Training	Limited lecturer preparedness; low self-efficacy reported (Gustilo et al., 2024)	Resource and time constraints; uneven motivation; lack of AI expertise	Provide structured professional development; offer discipline-specific case examples; build continuous support systems and communities of practice
Assessment Redesign	Emerging practices in early adoption stages (Kővári, 2025)	Increased workload; academic resistance; limited evidence of longterm outcomes	Prioritize authentic, process-based, and multimodal assessments; share evidence-based models; support gradual implementation
Equitable Access	Unequal AI access across socioeconomic and geographic groups (Costa et al., 2024)	Financial barriers to licenses; tech infrastructure gaps; digital literacy inequality	Institutional AI subscriptions; universal design approaches; targeted digital literacy and support programs.
Student AI Literacy	Early stage; few structured AI literacy initiatives (Turingan, 2025)	Crowded curricula; uneven lecturer preparedness; unclear competency standards	Embed AI literacy across courses; model ethical & critical AI use; evaluate AI skill development

Evaluation and Research	Limited longitudinal and cross-context evidence	Rapid change outpaces research; methodological complexity; resource demands.	Conduct longitudinal and comparative studies; enable cross-institutional research; use mixed methods and iterative evaluation.

Note: Evidence derived from included studies; table summarises reported strategies rather than evaluated effectiveness.

To avoid equating evidence of differing methodological strength, thematic interpretations were examined in relation to study design and quality appraisal. Pedagogical benefit claims were supported by a mix of experimental and quasi-experimental studies ($n = 5$) and self-report survey studies ($n = 8$), indicating stronger evidence for feedback and language support effects than for long-term learning outcomes. In contrast, many ethical and integrity concerns were derived from conceptual analyses, reviews, and cross-sectional surveys, highlighting perceived rather than experimentally verified risks. Inclusivity-related findings were primarily supported by targeted empirical studies involving students with disabilities, whereas institutional governance claims were largely grounded in policy analyses and synthesis studies. MMAT ratings were used to inform the weighting of evidence, with findings from higher-quality studies receiving greater interpretive emphasis.

5. Discussion

5.1 Overview of Key Findings

This review synthesised evidence from 27 studies on the use of AI-based writing (AIW) tools in higher education. Overall, the findings show a clear pattern: AIW tools offer practical benefits, but they also introduce important risks and challenges.

As shown in Table 4, strong evidence supports the role of AIW tools in improving writing efficiency and feedback, based on multiple high-quality studies.

Table 4. Summary of the Major Analytical Claims of the review, the number and type of supporting studies, and the relative strength of evidence based on MMAT appraisal

Major Claim	Supporting Studies (n)	Dominant Study Types	Evidence Strength
AIW tools improve writing efficiency and feedback	12	Experimental; mixed methods; surveys	High
AIW tools enhance grammar and language support	10	Experimental; classroom studies	High
AIW tools increase plagiarism and integrity risks	15	Surveys; reviews; conceptual	Medium
Over-reliance may erode critical thinking	11	Surveys; conceptual	Medium
AIW improves accessibility for disabled students	6	Targeted empirical studies	Moderate–High
Sustainable AIW use requires institutional governance	16	Reviews; policy analyses	Medium

The most consistently reported benefits include improved writing efficiency, faster feedback, and better language support, especially for non-native English speakers (e.g., Angelos et al., 2024; AlAfnan et al., 2023; Gustilo et al., 2024). At the same time, concerns about academic integrity, over-reliance, and reduced critical thinking appear across a large proportion of studies (e.g., Duah & McGivern, 2024; Kővári, 2025). Issues related to governance, policy, and sustainability are also widely discussed, while inclusivity and accessibility receive comparatively less attention.

Importantly, the strength of evidence varies across themes. Some findings are supported by multiple high-quality studies, while others rely on more limited or mixed evidence. This suggests that conclusions should be interpreted with appropriate caution.

Taken together, the results show that AIW tools are not inherently beneficial or harmful. Their impact depends on how they are used, the context in which they are applied, and the level of support provided to students.

5.2 Geographical and Disciplinary Variations

The review also reveals clear geographical differences in research focus and regulatory concerns. Studies conducted in North America and Europe tend to foreground academic integrity, AI detection, and assessment security, whereas research from Asia-Pacific and Middle Eastern contexts more often highlights pedagogical and equity-related benefits, particularly for language support and access (AISofi, 2024; Costa et al., 2024; Zhao et al., 2024).

These differences highlight the limitations of uniform governance approaches. In multilingual and under-resourced contexts, AI-assisted writing may function as a legitimate accessibility tool rather than primarily as a threat to academic integrity (Costa et al., 2024). By contrast, in highly resourced and monolingual settings, concerns about authorship attribution and assessment validity appear more pronounced (Kővári, 2025). Effective governance therefore requires context-sensitive policy frameworks rather than universal prohibition or uncritical adoption.

Disciplinary variation remains insufficiently examined. Most studies treat academic writing as a general skill, despite clear differences across STEM, humanities, and professional disciplines. This represents a significant gap, as the pedagogical implications of AIW use are likely to vary across disciplinary epistemologies and assessment traditions (Oates & Johnson, 2025).

5.3 Theoretical Interpretation: Why and When AIW Tools Work

The findings can be better understood through the perspectives of Vygotsky's Zone of Proximal Development (ZPD) and Zimmerman's self-regulated learning theory.

From a sociocultural perspective, AIW tools can act as mediating supports that assist learning within the learner's ZPD. Features such as automated feedback and writing suggestions can provide structured guidance that helps students refine their work beyond their current independent ability. However, this support is effective only when learners actively engage with it. When students accept AI-generated outputs without reflection, the tool bypasses the developmental processes central to learning.

Self-regulated learning theory further explains this variation. Students who use AI tools strategically, by evaluating, revising, and integrating suggestions, tend to benefit more (Zimmerman, 2002). In contrast, students who rely on AI passively are more likely to develop dependency and show reduced critical engagement.

This explains why the same tools produce different outcomes across contexts. Positive outcomes occur when AI supports cognitive processes, whereas negative outcomes emerge when it replaces them.

5.4 Balancing Efficiency and Learning

One of the most frequently reported benefits of AIW tools is efficiency. These tools can reduce the time required for drafting, editing, and language correction (Gustilo et al., 2024; Ely & Rad, 2024). This is particularly helpful for novice writers and multilingual learners.

However, efficiency should not be treated as an unquestioned advantage. Writing is not only a product but also a cognitive process that involves reflection, organization, and revision. These processes require effort and are essential for learning.

When AI tools reduce effort too much, they may also reduce learning. Students may produce better texts but engage less deeply with the content. This creates a tension between speed and understanding.

At the same time, efficiency can support learning when it reduces lower-level cognitive load. For example, grammar correction tools can free time for higher-order thinking. Therefore, the impact of efficiency depends on whether it supports or replaces meaningful engagement.

5.5 Differentiated Effects by AIW Tool Type

A key contribution of this review is the distinction between different types of AIW tools. These tools differ in function, and therefore in their pedagogical value and risks.

Grammar and style tools, such as Grammarly, operate on student-produced text and provide corrective feedback. These tools support revision and align with process-based writing pedagogy, particularly for multilingual learners (e.g., Zawacki-Richter et al., 2019).

Paraphrasing tools, such as QuillBot, can assist with language expression but may lead to superficial rewriting if used without understanding. This raises concerns about reduced comprehension and weak synthesis skills.

Generative AI systems, such as ChatGPT, represent a more fundamental shift. These tools can produce complete text outputs, potentially bypassing key stages of thinking and writing. While they may support brainstorming, they also raise concerns about authorship and academic integrity (e.g., Kasneci et al., 2023).

Overall, the evidence suggests that AIW tools should not be treated as a single category. Their educational impact depends on their function, and different tools require different pedagogical and policy approaches.

5.6 Equity, Power, and Structural Implications

AIW tools also raise important questions about equity and power. While they are often presented as tools that democratize access, their benefits are not equally distributed.

First, many tools are aligned with standardized academic English, which may disadvantage students with diverse linguistic backgrounds (Zawacki-Richter et al., 2019). This can reinforce existing norms rather than support linguistic diversity.

Second, access to advanced AI tools is often linked to cost. Paid features and subscription models may create inequalities between students with different financial resources (Zhao et al., 2024; Costa et al., 2024).

Third, AIW tools operate within commercial systems where user data may be collected and used. This raises concerns about privacy, ownership, and the commodification of student work (Kasneci et al., 2023).

These issues suggest that AIW integration is not only a pedagogical matter but also a social and institutional one. Addressing these concerns requires critical and inclusive approaches.

5.7 Sustainability Considerations

AIW tools are likely to remain part of higher education in the long term, which raises important sustainability questions.

From a social perspective, sustainability involves equitable access, responsible use, and long-term learning outcomes. Institutions need to ensure that all students can benefit from these tools.

Environmental sustainability is also an emerging concern. Large AI systems require substantial computational resources, contributing to energy use and carbon emissions (Ritchie, 2025). This issue has received limited attention in educational research but is increasingly relevant.

Institutions should therefore consider sustainability when adopting AI tools, including responsible procurement and usage practices.

5.8 Policy Implications

The findings of this review highlight the need for clear and actionable policies.

One of the key implications is that institutions should differentiate between tool types and develop appropriate guidelines for their use. Though grammar tools can be broadly permitted as they generally deal with language accuracy, generative AI tools require more careful regulation (Kasneci et al., 2023).

Ensuring transparency in AI use is also essential. Students should disclose how AI tools are used in their work. This would ensure academic integrity and promote ethical use of generative AI.

Assessment practices would also need to evolve. Process-based assessments, such as writing drafts and conducting reflections, can better capture student learning and help circumvent unethical and overuse of AI tools.

Furthermore, training is critical for both students and educators. Suitable training sessions should be conducted to provide guidance on responsible AI use.

Finally, given the rapid evolution of AI tools, policies should be regularly reviewed to keep pace with technological change.

5.9 Limitations

This review has several limitations. It includes only peer-reviewed studies, which may exclude relevant insights from grey literature. This may also introduce publication bias, as positive findings are more likely to be published.

In addition, the number of included studies is relatively small, and many rely on self-reported data. This might limit the strength of the conclusions.

5.10 Future Research Directions

Future research should examine long-term effects of AIW tools on writing development. More experimental and longitudinal studies are needed.

Research should also explore differences across student populations, disciplines, and contexts. In addition, more work is needed on institutional implementation and policy effectiveness.

6. Conclusion

AI-based writing (AIW) tools are rapidly reshaping academic writing practices in higher education. This review shows that these tools offer clear benefits, including improved writing efficiency, immediate feedback, and enhanced language support, particularly for multilingual learners. At the same time, they introduce important challenges related to academic integrity, over-reliance, and the potential weakening of critical thinking and writing development.

The findings highlight that the impact of AIW tools is not determined by the technology itself, but by how it is used within pedagogical and institutional contexts. When integrated thoughtfully, AIW tools can function as supportive learning aids that enhance student engagement and development. However, when used without guidance, they may replace essential cognitive processes and undermine meaningful learning.

This review also demonstrates that AIW tools should not be treated as a single category. Different tools serve different functions and therefore require differentiated pedagogical approaches and policy responses. In addition, issues of equity, access, and sustainability must be considered to ensure that the benefits of AI are distributed fairly and responsibly.

Overall, the challenge for higher education is not whether to adopt AIW tools, but how to integrate them in ways that support learning, uphold academic integrity, and promote inclusive and sustainable educational practices. Thoughtful design, clear policies, and ongoing evaluation will be essential to ensure that AI serves as a tool for learning rather than a substitute for it.

Funding

This research is funded by the Research Incentive Grant for Teaching and Learning (GIPP), Universiti Putra Malaysia, grant no. 9323840.

Acknowledgments

The authors would like to thank the anonymous reviewers of this article. The authors would also like to acknowledge the support of the university-driven research program on lifelong learning and reskilling.

Declaration

The authors would like to declare that AI tools were used solely for language editing and clarity enhancement. No AI tools were used for data extraction, thematic coding, quality appraisal, synthesis, or interpretation of findings. All analytical decisions were made by the authors.

Data Availability Statement

The data extraction spreadsheet and the coding framework used for the thematic synthesis are available as supplementary materials. To support transparency and replicability, anonymised versions of these materials have also been deposited in an open-access repository, which can be accessed at: **[DOI: 10.17605/OSF.IO/F8P6C]**.

References

- AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3(2). <https://doi.org/10.37965/jait.2023.0184>
- Almusharraf, M. I. N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*. <https://doi.org/10.30935/cedtech/13605>
- Angelos, C., Antonia, S., Athanasia, K., Maria, K., Anastasia, M., Marina, M., & Sofia, M. (2024). Impact of artificial intelligence and virtual reality on educational inclusion: A systematic review of technologies supporting students with disabilities. *Neveléstudomány*. <https://doi.org/10.3390/educsci14111223>
- Bui, T. T. U., & Tong, T. V. A. (2025). The impact of AI writing tools on academic integrity: Unveiling English-majored students' perceptions and practical solutions. *AsiaCALL Online Journal*, 16(1), 83–110.
- Carhuaz Valdez, D. D., & Arcata-Maquera, E. J. (2024). *University education in the era of generative artificial intelligence: Is academic integrity at risk?* *Ingeniería Investiga*, 6. <https://doi.org/10.47796/ing.v6i00.967>
- Charmaine, B. (2023). To use or not to use ChatGPT and assistive artificial intelligence tools in higher education institutions? The modern-day conundrum – students' and faculty's perspectives. *Equity in Education & Society*. <https://doi.org/10.1177/27526461231215083>
- Costa, K., Ntsobi, P. M., & Mofolo, L. (2024). Challenges, benefits and recommendations for using generative artificial intelligence in academic writing: A case of ChatGPT. *Medicon Engineering Themes*, 7(4).
- Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. A. Buchanan & A. Bryman (Eds.), *The SAGE handbook of organizational research methods*. Sage Publications.
- Duah, J. E., McGivern, P. J., & Technology, L. (2024). How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies. *International Journal of Information and Learning Technology*, 41(2), 180–193.
- Ely, S. E., & Rad, M. (2024). Examining the opportunities and challenges of using artificial intelligence for engineering technical writing courses. <https://doi.org/10.18260/1-2--47381>
- Flores, R. E. A., & Trujillo, S. M. (2025). Investigación científica e inteligencia artificial en estudiantes de posgrado: Un análisis cualitativo [Scientific research and artificial intelligence in graduate students. A qualitative analysis]. *European Public & Social Innovation Review*, 10, 1–17.
- Gallent, C.T., Zapata, A. G., & Ortego, J. L. H. (2023). El impacto de la inteligencia artificial generativa en educación superior: Una mirada desde la ética y la integridad académica. *RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa*. <https://doi.org/10.30827/relieve.v29i2.29134>
- Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in the AI era. *International Journal for Educational Integrity*, 20(1), Article 3. <https://doi.org/10.1007/s40979-024-00153-8>

- Hanlin, L., Yu, W., Siqi, L., & Changqing, H. (2024). The influence of GenAI on the effectiveness of argumentative writing in higher education: Evidence from a quasi-experimental study in China. *Journal of Asian Public Policy*. <https://doi.org/10.1080/17516234.2024.2363128>
- Hazem, Z., John, M., & Mike, K. (2023). Ethics of generative AI. *Journal of Medical Ethics*. <https://doi.org/10.1136/jme-2023-108909>
- Hind, A. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: A systematic review. *Arab World English Journal*. <https://doi.org/10.24093/awej/chatgpt.2>
- Imran, M., & Norah, A. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*. <https://doi.org/10.30935/cedtech/13605>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103. <https://doi.org/10.1016/j.lindif.2023.102274>.
- King, C., Letlhogonolo Nolo, M., & Mfanelo Patrick, N. (2024). Challenges, benefits and recommendations for using generative artificial intelligence in academic writing: A case of ChatGPT. *Medicon Engineering Themes*, 7(6), 3–38. <https://doi.org/10.31222/osf.io/7hr5v>
- Kleopatra, N. (2024). Generative artificial intelligence in higher education: Exploring ways of harnessing pedagogical practices with the assistance of ChatGPT. *International Journal of Curriculum and Education*. <https://doi.org/10.47852/bonviewijce42022489>
- Konstantinos, T. K. (2025). *The role of ChatGPT in academic writing: Pedagogical and ethical dimensions*. <https://doi.org/10.56294/ai2025420>
- Kovari, A. (2025). A systematic review of AI-powered collaborative learning in higher education: Trends and outcomes from the last decade. *Social Sciences & Humanities Open*, 11, 101335. <https://doi.org/10.1016/j.ssaho.2025.101335>
- Leon, S., Ye, C., Ying, H. G., & Zhe, C. (2025). *Finding your voice: Using generative AI to help international students improve writing*. <https://doi.org/10.20944/preprints202502.1524.v1>
- Mohammed Siddique, K. (2025). Written assignments and ethical considerations of artificial intelligence in higher education. *International Journal of Linguistics, Literature and Translation*, 8(9), Article 27. <https://doi.org/10.32996/ijllt.2025.8.9.27>
- Mokhine-Martins, M., Deka, J., & Costa, K. (2024). Positive psychology interventions and initiatives – The Ubuntu approach – A mapping review protocol. *Journal of Research Methodologists*, 3(1). <https://doi.org/10.63591/GCAR.JRM.3.1.76>
- Moorhouse, B. L., Wan, Y., Wu, C., Wu, M., & Ho, T. Y. (2025). Generative AI tools and empowerment in L2 academic writing. *System*. <https://doi.org/10.1016/j.system.2025.103779>
- Muhammad, A., Aisha, A., Sadia, Q., Muhammad Sabboor, H., & Sadia, B. (2025). Academic writing issues in the AI era: Opportunities and challenges for applied and social sciences in Pakistan. *Academic Journal*, 4(3). <https://doi.org/10.63056/acad.004.03.0583>

Oates, A., & Johnson, D. (2025). ChatGPT in the classroom: Evaluating its role in fostering critical evaluation skills. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-024-00452-8>

Promethi Das, D., & Yixin, C. (2025). The role of AI in academic writing: Impacts on writing skills, critical thinking, and integrity in higher education. *Societies*. <https://doi.org/10.3390/soc15090247>

Quynh Hoa, N. (2023). AI and plagiarism: Opinion from teachers, administrators and policymakers. *PAIC*. <https://doi.org/10.54855/paic.2346>

Romaric, D. (2024). ChatGPT has eaten my assignment: A student-centric experiment on supervising writing processes in the AI era. *EDUCON Proceedings*. <https://doi.org/10.1109/educon60312.2024.10578917>

Rosario Michel-Villarreal, E. L., Vilalta-Perdomo, E. L., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Flor Silvestre, G. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Neveléstudomány*. <https://doi.org/10.3390/educsci13090856>

Saeed Awadh, B.-N., Mouad, S., & Mohamed, B. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*. <https://doi.org/10.1016/j.techsoc.2023.102370>

Santosh, K., Patil, V. D., Parikshit, N. M., & Vishal, M. (2023). Concept and administration for developing special education using AI. In *Advances in Educational Technologies and Instructional Design* (Chapter 11). <https://doi.org/10.4018/979-8-3693-0378-8.ch011>

Sarin, S., & Kimkong, H. (2024). Opportunities, challenges, and strategies for using ChatGPT in higher education: A literature review. *Journal of Digital Educational Technology*. <https://doi.org/10.30935/jdet/14027>

Stefan, K., & Sebastian, W. (2023). Navigating the impact of ChatGPT/GPT-4 on legal academic examinations: Challenges, opportunities, and recommendations. *Media Iuris*. <https://doi.org/10.20473/mi.v6i2.45270>

Thaweesak, C. (2025). Generative artificial intelligence in academic writing in higher education: A systematic review. *Edelweiss Applied Science and Technology*. <https://doi.org/10.55214/25768484.v9i4.6128>

Ting, W., Brady, D. L., Agostino, M., Pagano, A., Zoë, T., & Jenny, P. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences*. <https://doi.org/10.3390/app13116716>

Tom, F., & Nick, B. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Neveléstudomány*. <https://doi.org/10.3390/educsci13111109>

Tony, B., Cristóbal, C., Olga, M., & Steve, W. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-020-00218-x>

Turingan, V. T. (2025). Exploring students' perspectives on utilization of artificial intelligence (AI) writing tools through sequential explanatory mixed method study. *International Journal of English Literature and Social Sciences*, 10(4). <https://doi.org/10.22161/ijels.104.37>

Wang, Y. D., Chen, J., Yuxin, W., Xi, Y., & Lixia, H. (2025). Impact of generative artificial intelligence on scientific paper writing and regulatory pathways. *Editing Practice*. <https://doi.org/10.54844/ep2025.1011>

Wojciech, W. (2023). ChatGPT in higher education. *International Journal of New Economics and Social Sciences*. <https://doi.org/10.5604/01.3001.0053.9601>

Xhulio, M., & Muhamet, Z. (2024). Using AI to improve accessibility and inclusivity in higher education for students with disabilities. <https://doi.org/10.1109/ithet61869.2024.10837607>

Xianghan, O. D. (2024). Generative AI: Is it a paradigm shift for higher education? *Studies in Higher Education*. <https://doi.org/10.1080/03075079.2024.2332944>

Xiaoli, X., & Sheng, D. (2023). Opportunities, challenges, strategies, and reforms for ChatGPT in higher education. <https://doi.org/10.1109/eki61071.2023.00010>

Yanhua, L., Jaek, P., & Sean, M. (2024). Using generative artificial intelligence/ChatGPT for academic communication: Students' perspectives. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12574>

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16. <https://doi.org/10.1186/s41239-019-0171-0>

Self-Efficacy among Higher Vocational Education College Educators: A Systematic Literature Review

ABSTRACT

Authors: Shi Yinging, Muhd Khaizer Omar* & Normala Ismail
Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:

khaizer@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Teacher self-efficacy is a core topic in psychology and education, but relevant research in higher vocational education remains scarce, leaving a significant research gap. Literature focusing on vocational education teachers is particularly limited. **Methods:** This paper uses a systematic literature review to comprehensively review the literature on teacher self-efficacy in higher vocational education, focusing on conceptualization, research content, cultural context, measurement tools, research methods (quantitative, qualitative, intervention, and mixed methods), and study types (cross-sectional, longitudinal, and intervention studies). Furthermore, we summarize key themes and core findings of each study and analyze publication year and research location to explore the temporal and spatial distribution of this field. **Conclusion:** The 20 studies included reveal that most research on teacher self-efficacy focuses on teaching contexts, particularly in specialized fields such as technical subjects and writing self-efficacy. These studies use diverse methods and tools but lack systematicity and depth. Asia is the primary focus of research, where the cultural context reflects regional characteristics. Empirical research in higher vocational education remains limited, and future studies should explore cross-cultural variations to support targeted professional development and improve vocational education quality.

Keywords: teacher self-efficacy, higher vocational education, TVET, systematic literature review

1. Introduction

Teacher-related psychological resources play a crucial role in shaping the effectiveness of higher vocational education (HVE), particularly in contexts characterized by rapid technological change, diverse student needs, and strong alignment with labor market demands (OECD, 2021; UNESCO, 2022). Among these resources, teacher self-efficacy (TSE), grounded in Albert Bandura's social cognitive theory, has been widely recognized as a central construct influencing teachers' instructional practices, persistence, and adaptability (Bandura, 1997; Tschannen-Moran & Hoy, 2001). TSE refers to teachers' beliefs in their capability to organize and execute the actions required to accomplish specific teaching tasks (Bandura, 1997). Empirical studies have consistently shown that higher levels of TSE are associated with improved teaching effectiveness, greater student engagement, and enhanced learning outcomes (Holzberger et al., 2013; Klassen & Tze, 2014). In vocational education settings, where teaching often involves integrating theoretical knowledge with practical skills and responding to industry-oriented expectations, strong self-efficacy beliefs are particularly important for maintaining teaching quality and fostering student learning outcomes (Cai & Lin, 2021; OECD, 2021).

Despite the well-established importance of TSE in general education research, its application in higher vocational education remains relatively underexplored and fragmented (Klassen et al., 2011; Zee & Koomen, 2016). Existing studies in HVE contexts have primarily focused on specific domains such as technology integration, academic writing, or subject-specific teaching efficacy, rather than providing a comprehensive understanding of TSE as a multidimensional construct (Fackler & Malmberg, 2016; Perera et al., 2019). Moreover, several unresolved issues persist in the literature. First, there is considerable variation in how TSE is conceptualized and measured, with studies employing both general scales (e.g., Teacher Self-Efficacy Scale) and task-specific instruments, leading to inconsistencies in findings (Tschannen-Moran & Hoy, 2001; Klassen et al., 2011). Second, most studies adopt cross-sectional designs, limiting insights into the dynamic development of self-efficacy over time (Zee & Koomen, 2016). Third, research is geographically concentrated in certain regions—particularly in Asia—raising concerns about the cross-cultural generalizability of results (Klassen et al., 2009; Vieluf et al., 2013). Finally, the unique contextual features of vocational education, such as its strong emphasis on practice-oriented teaching and industry collaboration, have not been sufficiently theorized in relation to TSE, leaving important context-specific mechanisms underexplored (Lucas et al., 2012; Wheelahan, 2015).

Although previous reviews have examined teacher self-efficacy in broader educational contexts, few have focused specifically on higher vocational education, and even fewer have systematically analyzed the conceptualization, measurement approaches, and contextual characteristics of TSE within this domain (Zee & Koomen, 2016; Klassen et al., 2011). As a result, the current body of knowledge lacks an integrated and critical synthesis that can clarify research trends, methodological patterns, and key gaps in the field. To address this limitation, the present study conducts a systematic literature review of empirical research on teacher self-efficacy in higher vocational education, following established guidelines for systematic reviews (David Moher et al., 2009; Page et al., 2021). By synthesizing evidence across studies, this review aims to provide a comprehensive overview of how TSE is conceptualized and operationalized, identify major research themes and influencing factors, examine methodological characteristics, and highlight directions for future research. In doing so, the study contributes to a more coherent understanding of TSE in vocational education and offers evidence-based insights for enhancing teacher development and educational quality in HVE contexts.

1.1. Teacher Self-efficacy

Self-efficacy is a core construct in Bandura's social cognitive theory, defined as "the belief in one's ability to organize and execute a course of action necessary to achieve achievement" (Bandura, 1997). Teacher self-efficacy (TSE) is a specific concept in educational contexts, understood as a teacher's belief in their ability to complete all tasks necessary to promote effective teaching (Holzberger et al., 2013; Tschannen-Moran et al., 1998). In disciplines such as mathematics and science teaching, TSE has been

extensively researched for decades and is even considered "nearly mature" (Tschannen-Moran et al., 1998). In contrast, research in this area of vocational education has lagged significantly, and systematic discussion remains limited. Na and Isa (2024) emphasize the critical importance of TSE in higher vocational education; they found that teachers with higher self-efficacy tend to achieve higher-quality teaching outcomes, which in turn directly improves student performance. This conclusion is consistent with their research on the relationship between self-efficacy and teaching quality. Yang et al. (2024) further suggested that improving teachers' knowledge, particularly when teaching sustainable development-related content in higher vocational education, can positively promote their self-efficacy. This finding not only expands the connection between self-efficacy and teaching preparation but also provides new directions for future research. Furthermore, Kholifah et al. (2023) confirmed the close relationship between self-efficacy and teaching quality among vocational teachers by linking psychological well-being with teaching performance. Andersson et al. called for further exploration of the multiple factors influencing TSE, particularly in the context of professional development within the vocational education context. Overall, the existing literature not only emphasizes the importance of self-efficacy in higher vocational education but also reveals significant research gaps in this area. A comprehensive study of the unique factors influencing vocational education teachers' self-efficacy will facilitate the development of more targeted professional development strategies and improve student learning outcomes in vocational courses.

2. Theoretical and Conceptual Framework

This study is grounded in the social cognitive theory proposed by Albert Bandura (1997), which conceptualizes self-efficacy as individuals' beliefs in their capabilities to organize and execute the actions required to achieve specific outcomes. Within this framework, teacher self-efficacy (TSE) is understood as teachers' beliefs in their ability to effectively perform teaching-related tasks. Social cognitive theory emphasizes the dynamic interaction among personal, behavioral, and environmental factors—referred to as reciprocal determinism—suggesting that self-efficacy is continuously shaped through experience and context rather than being a fixed personal trait.

According to this theory, self-efficacy beliefs are developed through four primary sources: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional states (Bandura, 1997). Mastery experiences, which refer to individuals' direct experiences of success, are considered the most influential source. In higher vocational education (HVE), successful instructional practices, effective classroom management, and positive student outcomes contribute to strengthening teachers' efficacy beliefs. Vicarious experiences arise from observing others' performance, such as through peer collaboration, classroom observation, and professional training, which are particularly relevant in practice-oriented teaching environments. Social persuasion includes feedback and support from school leaders, colleagues, and institutional structures, which can enhance teachers' confidence in their instructional abilities. In addition, physiological and emotional states, such as stress or emotional well-being, influence how teachers interpret their competence and cope with teaching demands.

In the context of higher vocational education, teacher self-efficacy demonstrates strong contextual specificity. Compared with general education, vocational education places greater emphasis on practice-oriented instruction, the integration of theoretical knowledge with technical skills, and alignment with labor market demands. These characteristics require teachers to adapt to rapidly evolving technologies, engage in hands-on teaching, and address diverse and task-specific instructional challenges. As a result, TSE in HVE is inherently multidimensional, encompassing not only general teaching efficacy but also efficacy beliefs related to specific tasks and technology-enhanced teaching contexts. This multidimensionality reflects the varied demands placed on teachers across different instructional situations.

Based on this theoretical foundation, the present study adopts social cognitive theory as a guiding lens to synthesize existing research on teacher self-efficacy in higher vocational education. This framework provides a coherent basis for interpreting how TSE is conceptualized, what factors influence its

development, and how it manifests across different contexts. By grounding the review in a well-established theoretical perspective, the study enhances the overall coherence and explanatory strength of the literature synthesis.

3. Objectives

This systematic literature review aims to comprehensively review the current status of research on teacher self-efficacy in the field of vocational education. Specific objectives include: (1) analyzing the temporal and regional distribution characteristics of relevant literature; (2) comprehensively summarizing the definition and conceptual connotation of teacher self-efficacy in the context of vocational education; (3) identifying and summarizing the measurement tools used in various studies; (4) sorting out the topics related to teacher self-efficacy and their influencing factors; and (5) integrating the main empirical findings of the selected literature. The specific research questions are as follows:

1. RQ1: How is TSE conceptualised and measured in HVE research?
2. RQ2: What factors influence TSE in HVE contexts?
3. RQ3: What methodological patterns exist in TSE research?
4. RQ4: What research gaps and future directions emerge?

4. Prior Review Evidence

Although teacher self-efficacy (TSE) has received extensive attention in educational research, existing review studies have primarily focused on general education contexts, particularly in primary and secondary education. For instance, prior reviews have systematically examined the conceptualization, measurement, and outcomes of TSE, highlighting its significant role in enhancing teaching effectiveness, promoting student engagement, and improving instructional practices (Klassen et al., 2011; Zee & Koomen, 2016).

However, systematic reviews that specifically address teacher self-efficacy in higher vocational education (HVE) remain limited. Existing studies in vocational education tend to be fragmented, often focusing on specific aspects such as technological self-efficacy or task-specific teaching efficacy, rather than providing a comprehensive synthesis of TSE within this context (Watt & Richardson, 2008; Fives & Gill, 2015). As a result, there is still a lack of systematic understanding regarding the conceptualization, measurement, and underlying mechanisms of TSE in HVE settings.

In addition, previous review studies have identified several ongoing challenges in TSE research, including inconsistencies in measurement approaches, the dominance of cross-sectional designs, and insufficient attention to contextual and cultural variations (Klassen et al., 2011; Zee & Koomen, 2016). These issues are particularly salient in vocational education, where teaching is closely linked to practice-based instruction, industry collaboration, and skill-oriented training, making its educational context distinct from that of general education (Fuller & Unwin, 2011).

In response to these gaps, the present study conducts a systematic literature review focusing specifically on teacher self-efficacy in higher vocational education. By synthesizing empirical studies in this domain, this review aims to provide a more comprehensive and context-sensitive understanding of TSE in HVE, thereby extending existing review literature and contributing to the advancement of vocational education research.

4. Methodology

This study employed a systematic literature review (SLR) approach. A systematic review is considered a research strategy for identifying and synthesizing existing research evidence of sufficient quality on a specific topic (Victor, 2008). To answer our research questions, we searched and evaluated preliminary data from databases relevant to our research topic. During the literature selection and analysis process,

we strictly adhered to the PRISMA checklist (Page et al., 2021). The PRISMA flow chart (Figure 1) illustrates the various stages of this systematic review, including the number of references initially identified, the number of excluded studies, and the final number of studies included in the analysis.

4.1 Search Strategy

We searched for literature through Scopus, Web of Science, and Google Scholar electronic databases. The starting year of 2010 was selected to capture research conducted in the context of recent developments in vocational education, including increased emphasis on skills-based training and internationalization. The purpose of this study was to summarize the psychological problems and influencing factors of teachers in the field of vocational education by including articles on the self-efficacy of vocational education teachers. Boolean logic operators AND or OR were used to connect keywords. The specific main keywords of this study include (1) teacher self-efficacy and (2) vocational education. The specific research strategy is as follows:

("self-efficacy" OR "teacher efficacy" OR "teaching efficacy") AND ("vocational education" OR "technical education" OR "vocational and technical education" OR "VET" OR "TVET" OR "career and technical education" OR "vocational teachers" OR "technical teachers" OR "vocational school teachers")

4.1.1. Scopus

The search was conducted in August 2025 from Mainland China, using institutional access to Scopus, on 12 August 2025 (China Standard Time, UTC+8). The following fields were exported for subsequent screening and analysis: authors, publication year, article title, abstract, author keywords, source title, DOI, author affiliations, and cited references.

The Scopus database was searched using the following exact search string:

TITLE-ABS-KEY ("self-efficacy" OR "teacher efficacy" OR "teaching efficacy") AND TITLE-ABS-KEY ("vocational education" OR "technical education" OR "vocational and technical education" OR "VET" OR "TVET" OR "career and technical education" OR "vocational teachers" OR "technical teachers" OR "vocational school teachers")

4.1.2. Web of Science

The search was performed in August 2025 from Mainland China, using institutional access, on 12 August 2025 (China Standard Time, UTC+8). Records were exported with full records and cited references, including authors, publication year, article title, abstract, author keywords, source title, DOI, author affiliations, and cited references, for subsequent screening and analysis.

The Web of Science Core Collection was searched using the following exact query:

TS = ("self-efficacy" OR "teacher efficacy" OR "teaching efficacy") AND TS = ("vocational education" OR "technical education" OR "vocational and technical education" OR "VET" OR "TVET" OR "career and technical education" OR "vocational teachers" OR "technical teachers" OR "vocational school teachers")

4.1.3. Google Scholar

Google Scholar was searched as a supplementary source to identify relevant studies not indexed in Scopus or Web of Science. Searches were conducted using Publish or Perish, which enables structured retrieval of Google Scholar records with a maximum export limit of 1,000 records per query. All Google Scholar records were merged with results from Scopus and Web of Science, and duplicate records were identified and removed using Zotero based on title, author, and publication year. During the screening

process, potentially relevant records, including those not indexed in major databases, were manually reviewed to ensure that no high-value studies were overlooked.

The search was conducted in August 2025 from Mainland China, on 13 August 2025 at (China Standard Time, UTC+8). Google Scholar results are ranked by relevance by default; therefore, all 1,000 retrievable records were exported and screened. Exported fields included authors, publication year, title, source, abstract, and citation count. All Google Scholar records were merged with Scopus and Web of Science results prior to deduplication. Duplicate records were identified and removed using Zotero, based on title, author, and publication year.

The following search terms were applied:

"self-efficacy" OR "teacher efficacy" OR "teaching efficacy"
AND
"vocational education" OR "technical education" OR "vocational and technical education" OR "VET"
OR "TVET" OR "career and technical education" OR "vocational teachers" OR "technical teachers"
OR "vocational school teachers"

4.2. Inclusion and exclusion criteria

This study established clear screening criteria based on the research objectives and conducted a systematic literature review using inclusion and exclusion criteria.

Inclusion criteria are as follows:

1. Articles must contain the following core keywords: teacher self-efficacy, self-efficacy, vocational education;
2. The publication language must be English;
3. Research methods are not limited to quantitative research, qualitative research, or mixed research, and all are included in the analysis.
4. Studies published between 2010 and 2025.

Exclusion criteria are as follows:

1. Literature whose research subjects are not vocational education teachers (e.g., research on primary and secondary school teachers is not included);
2. Literature whose research topics are not related to vocational education or teacher self-efficacy;
3. Non-empirical research, such as reviews, comments, editorials, and theoretical discussions;
4. Research materials for which the full text is unavailable.

4.3. Data extraction

During the identification phase, the researchers conducted a systematic search across multiple databases, identifying 2,069 records, including 647 from Web of Science, 422 from Scopus, and 1,000 from Google Scholar. Before formal screening, 190 duplicate records were removed, leaving 1,879 for preliminary screening. Subsequently, a preliminary review of titles and abstracts excluded 1,835 articles that did not meet the inclusion criteria, primarily because they were unrelated to the research topic, focused on non-educational contexts, targeted different populations, or did not address the core variables of interest. 44 full-text articles were retained, of which two were excluded due to unavailable original texts. Detailed reasons for study exclusion are reported in Supplementary Table S1. Ultimately, 42 articles entered the full-text evaluation phase. Following strict inclusion and exclusion criteria, 22 articles were further excluded for the following reasons: seven articles were not in the field of vocational education, two articles focused solely on students rather than teachers, three were review articles, and 10 were informal publications such as books, conference papers, and reports. A total of 20 eligible articles were included for subsequent systematic analysis and information extraction. The screening

process is detailed in Figure 1 (PRISMA flow chart). Document management and data processing were completed using Microsoft Office Excel and Zotero, respectively. Detailed findings and quality assessment are provided in Supplementary Table S2.

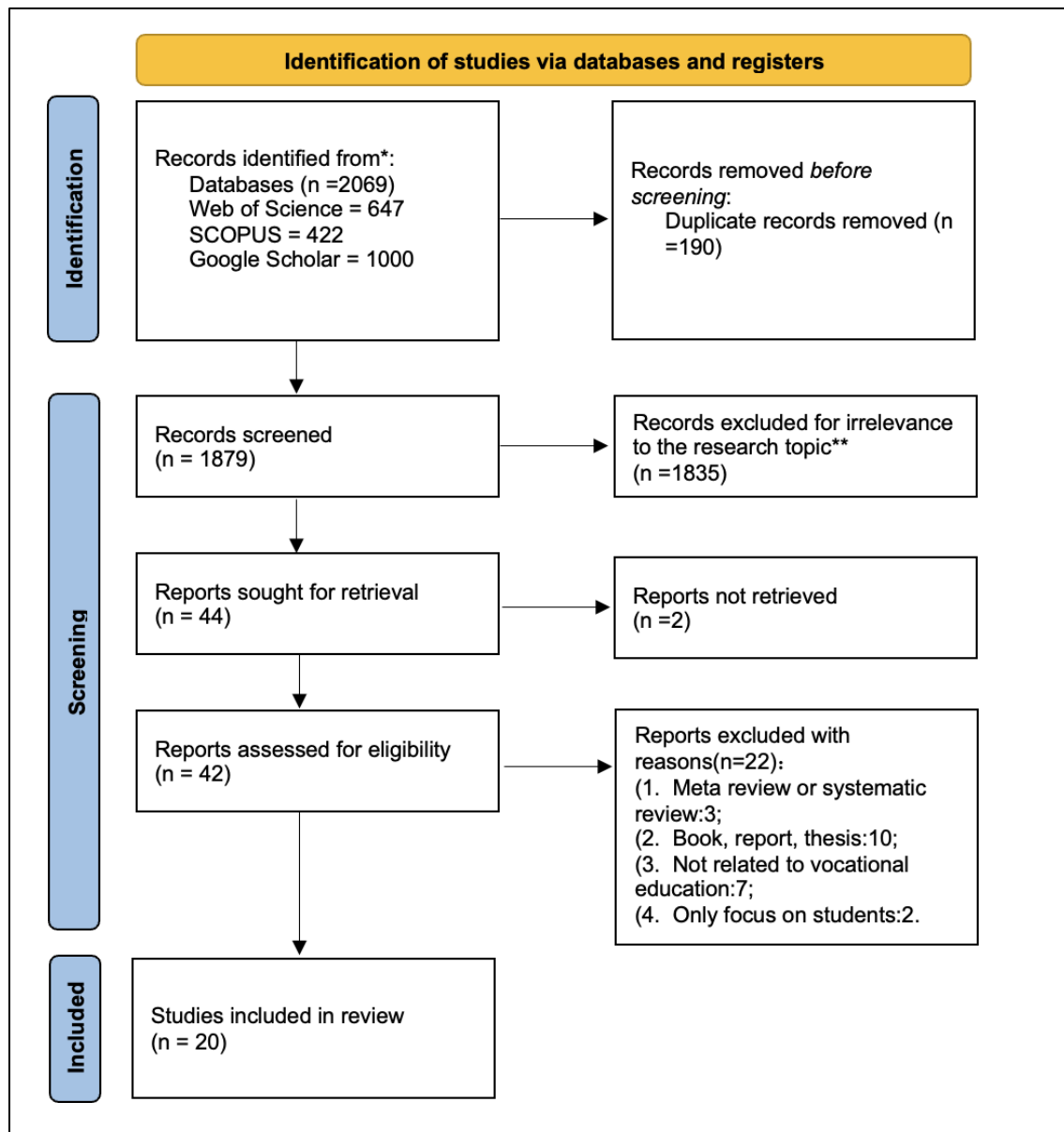


Figure 1. PRISMA Flow Diagram

4.4. Quality Assessment

To evaluate the methodological quality of the included studies, the Crowe Critical Appraisal Tool (CCAT) was utilized. The CCAT is suitable for appraising various research designs—including mixed-methods, quantitative, and qualitative approaches—making it an appropriate choice for the diverse methodologies present in the selected literature. Each study was assessed across eight domains, with scores assigned on a 0–5 scale for each domain. Scores for each domain were summed to produce a total score ranging from 0 to 40, which was then classified into three quality categories: High (≥ 30), Moderate (20–29), and Low (< 20). Detailed CCAT domain scores for all included studies are provided in Supplementary Table S3. Study quality was assessed using the CCAT and informed the narrative synthesis and interpretation of findings. To ensure the reliability of the study selection process, two reviewers independently screened the titles, abstracts, and full texts of all retrieved studies based on predefined inclusion and exclusion criteria. The level of agreement between the two reviewers was

high (Above 85%) throughout the screening process. Any discrepancies were discussed and resolved through consultation with a third reviewer to reach a final consensus. Although no statistical weighting or quantitative sensitivity analysis was conducted, greater interpretive emphasis was placed on evidence from studies with higher methodological quality. Key themes and conclusions were examined for consistency across studies of varying quality to ensure the robustness of the synthesis. This structured and transparent approach ensured consistent evaluation of methodological rigor, thereby enhancing the reliability and validity of the review's conclusions. The CCAT scores and attributes for each of the 20 studies are presented in Table 1. Specifically, studies with higher CCAT scores were given greater interpretive emphasis in the synthesis, while findings from studies with relatively lower quality were interpreted with caution. In addition, consistency of key findings across studies with different quality levels was considered to enhance the robustness of the conclusions.

4. Results

4.1. Study Selection and Characteristics

Research on self-efficacy is quite mature, and the literature on teacher self-efficacy is also reflected in multiple research fields. This study focuses on teacher self-efficacy in vocational and technical education. Figure 2 shows the publication trends of related research from 2010 to 2025. In 2010, there was only one related publication, and the annual number of publications remained low from then until 2022. However, the number of related studies increased significantly in 2023, reaching six, and in 2024 and 2025, four and three articles, respectively. The overall trend shows that academic attention to this topic has increased significantly in recent years.

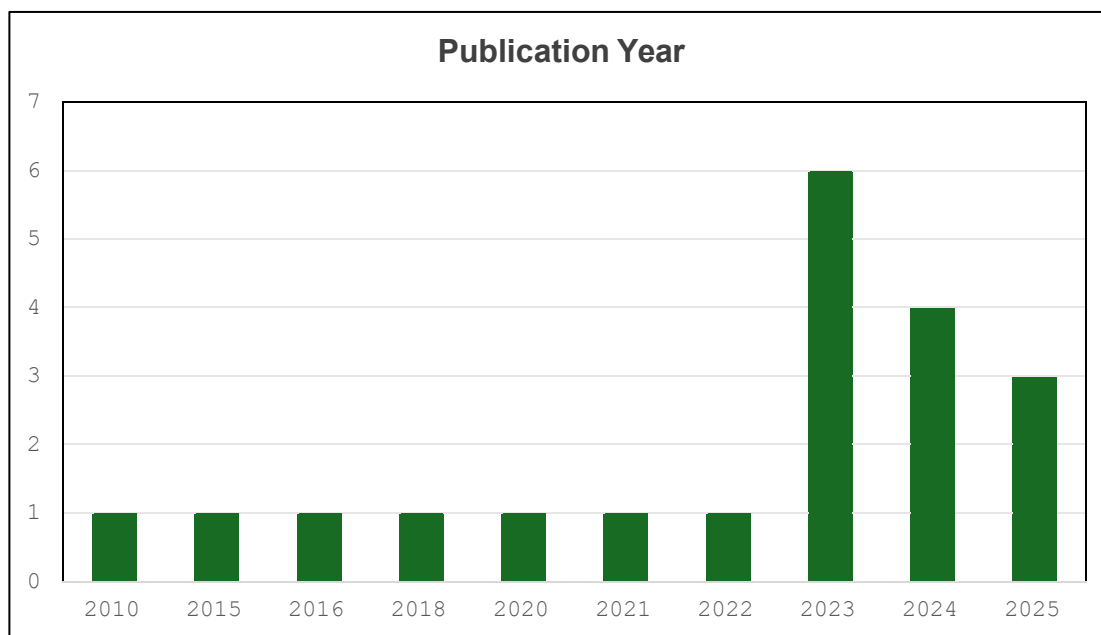


Figure 2. Distribution of Included Studies by Year of Publication

The pie chart in Figure 3 shows the distribution of publications on vocational education teacher self-efficacy by country. The results show that Asia dominates this research, contributing 64% of the research. Specifically, Indonesia leads with 32%, followed by China with 16%, Malaysia with 11%, and Thailand with 5%. Research from other regions primarily comes from the Netherlands (16%), Canada (10%), Switzerland (5%), and Nigeria (5%). This distribution highlights Asia's significant influence on this topic and reflects the field's internationalization.

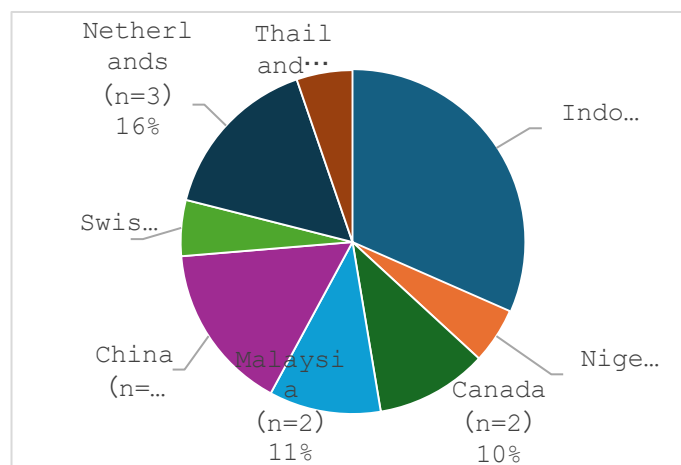


Figure 3. Regional Distribution of Vocational Education Publications

4.2. Conceptual Category

4.2.1. Definition of Teacher Self-efficacy

Across the reviewed literature, definitions of teacher self-efficacy consistently describe it as teachers' belief in their capability to successfully perform instructional tasks and influence student engagement and learning outcomes. Rooted predominantly in Bandura's self-efficacy theory (1997), these definitions highlight both the confidence and the judgment teachers hold regarding their teaching abilities in specific contexts, such as subject-specific instruction, vocational education, integration of educational technology, academic writing, and classroom management. Teacher self-efficacy is often conceptualized as multidimensional, encompassing domains such as instructional strategies, student engagement, and classroom management, and is closely linked to professional identity, resilience in the face of challenges, and the capacity to adopt innovative practices. This conceptual foundation underscores the importance of teacher self-efficacy as a critical determinant of teaching effectiveness and educational innovation.

Table 1. Definition of Teacher Self-efficacy Included in the Study

No.	Author(s)	Definition	CCAT Score
1	Agustina et al., 2020	Teachers believe they can influence student engagement and learning outcomes, especially when they are free from external pressure and possess a strong commitment to teaching (Barni et al., 2019).	35
2	Chou et al., 2010	Teachers' confidence and judgment of their ability to use computers to complete tasks in the teaching environment are part of computer self-efficacy, based on Bandura's self-efficacy theory (Compeau & Higgins, 1995; Wilfong, 2006).	36
3	Adebagbo, 2025	An individual's belief in their ability to complete a specific task (Bandura, 1997).	37
4	Gagnon & Dubeau, 2023	An individual's judgment and belief about whether they possess the necessary action capabilities to achieve a specific goal (Bandura, 1977).	38
5	Yang et al., 2024	Teachers believe they can effectively plan and implement teaching objectives in a specific subject (Bandura, 1994).	35
6	Li, 2025	Teachers' belief in their ability to influence student learning outcomes is a crucial component of their professional identity (Sun et al., 2025).	34
7	Mulyono et al., 2023	Teachers' perceptions of and confidence in their own abilities during academic writing and publication (Mulyono et al., 2023).	35
8	Anwar et al., 2024	Teachers' beliefs about their ability to influence student learning through instructional activities (Anwar et al., 2024).	38
9	Ismail et al., 2025	Teachers' beliefs about their ability to successfully adopt educational technology, implement innovative teaching practices, and enhance	37

		student engagement (Prasetya et al., 2021).	
10	Lauermann & Berger, 2020	Teachers' judgments about their ability to effectively promote engagement and learning outcomes, even with struggling students (Tschannen-Moran & Woolfolk Hoy, 2001).	37
11	Jaedun et al., 2022	Teachers' confidence in their ability to complete specific instructional tasks (Jaedun et al., 2022).	38
12	Beverborg et al., 2015	Teachers' beliefs about their ability to successfully perform tasks in instructional activities stem from factors such as mastery and experience (Bandura, 1997).	38
13	Runhaar et al., 2016	Teachers' belief in their ability to cope with various challenges in their professional environment (Runhaar et al., 2016).	37
14	Preechawong et al., 2021	Teachers' beliefs about their ability to design instruction, manage the classroom, and interact with students (Bandura, 1994).	36
15	Klaeijsen et al., 2018	Teachers' judgments about their ability to effectively handle new tasks or future challenges are distinct from emotional experiences of competence (Bandura, 1982).	35
16	Code et al., 2023	Teachers' beliefs about their ability to influence student learning and sustain teaching in higher vocational education settings (Findeisen et al., 2022).	34
17	Fang & Qi, 2023	Teachers' confidence in their ability to influence teaching and student learning (Fang & Qi, 2023).	38
18	Siswanto et al., 2023	Teachers' confidence in their effectiveness in student engagement, teaching strategies, and classroom management in face-to-face and online teaching (Siswanto et al., 2023).	39
19	Kholifah et al., 2023	Teachers' beliefs about their ability to organize and execute teaching activities to achieve specific learning objectives (Bandura, 1994).	36
20	Jin et al., 2021	Teachers' confidence in effectively executing teaching tasks emphasizes its importance in knowledge management and educational innovation (Tschannen-Moran & Hoy, 2001).	34

4.2.2. Measurement Tool of Teacher Self-efficacy

Across the reviewed studies, a range of instruments was employed to assess teacher self-efficacy, reflecting both general conceptualizations and context-specific applications. Table 2. shows the specific content. General teacher self-efficacy was commonly measured using Bandura's Teacher Self-Efficacy Scale and the Teacher Self-Efficacy Scale (TSES), which evaluate multiple domains such as instructional strategies, classroom management, and student engagement. In technology-integrated contexts, instruments such as the Computer Self-Efficacy Scale, Technological Self-Efficacy, and the Teachers' AI Self-Efficacy Questionnaire (TASQ) were used to capture teachers' confidence in adopting and applying digital tools and artificial intelligence in educational settings. Task-specific instruments, including the Academic Creativity and Self-Efficacy Scale (ACSES) and the Situated Academic Writing Self-Efficacy Scale (SAWSSES), focused on specialized competencies such as academic innovation and scholarly writing. Additionally, Occupational Self-Efficacy scales were applied to examine teachers' perceived capability to meet professional demands and adapt to workplace challenges. This diversity of measurement tools illustrates the multidimensional and context-dependent nature of teacher self-efficacy in vocational and technical education.

Table 2. Teacher Self-efficacy Measurement

Variables	Measurement
Teacher self-efficacy	Bandura's Teacher Self-Efficacy Scale
	Computer Self-Efficacy Scale
	Teachers' AI Self-efficacy Questionnaire (TASQ)
	Teacher Self-Efficacy Scale(TSES)
	Academic Creativity and Self-Efficacy scale (ACSES)
	Situated Academic Writing Self-Efficacy Scale (SAWSSES)
	Technological self-efficacy
	Occupational self-efficacy

4.2.3. Sub-category and Main Findings of Teacher Self-efficacy

An analysis of 20 included studies revealed that teacher self-efficacy (TSE) in vocational and technical education is closely related to multiple themes, exhibiting multidimensional and context-dependent characteristics. A thematic synthesis approach was used to organize the findings of the included studies. The themes were generated inductively based on the research focus and key variables reported in the literature, rather than through formal coding procedures. This approach allowed for a structured yet flexible synthesis of the existing evidence. Overall, these themes can be categorized into four main categories: professional practice and engagement. TSE is often associated with reflective practice, professional engagement, teaching effectiveness, and teaching quality, demonstrating its crucial role in improving teacher performance and maintaining motivation across diverse contexts, including face-to-face and online teaching. Second, organizational and contextual factors. Numerous studies have linked TSE to organizational climate, school leadership support, mentoring experience, and perceived organizational and innovative climates, demonstrating that supportive systems and environments contribute to enhancing teacher self-efficacy. Third, psychological and motivational dimensions. TSE is closely related to psychological capital, intrinsic motivation, professional adaptability, job satisfaction, and work engagement, highlighting the importance of psychological resources and motivational factors in supporting teachers' professional confidence. Fourth, the integration of technological competence and digitalization. Within the context of digitalization, TSE is closely linked to topics such as computer self-efficacy, technology acceptance, digital leadership, digital technology familiarity, and the continued use of e-learning, indicating that technology-related self-efficacy is becoming an integral part of the role of vocational college teachers. This thematic distribution (see word cloud, Figure 4) shows that TSE is not only related to teachers' professional identity and career development, but is also closely intertwined with organizational support and technology integration, highlighting its central role in enhancing the effectiveness and adaptability of vocational college teachers.

Comprehensive results from the 20 included studies indicate that teacher self-efficacy in vocational and technical education plays a central role in multiple aspects, including teaching, technology application, organizational support, and professional development. It not only directly influences teachers' teaching effectiveness, student learning outcomes, and innovative teaching behaviors, but also indirectly promotes teachers' professional practice and continuous development through mediating or moderating effects on factors such as principal digital leadership, organizational innovation climate, work engagement, professional motivation, and psychological well-being. Research indicates that teacher self-efficacy is influenced by a variety of factors, including mastery experience, vicarious experience, social persuasion, emotional state, career goal orientation, task dependency, mentorship, peer relationships, and support from school leadership. Furthermore, technological self-efficacy plays a key role in digital teaching and the sustained use of technology, enhancing teachers' innovative abilities and classroom teaching quality. Overall, teacher self-efficacy is not only a crucial psychological resource for teachers' professional competence but also a crucial bridge connecting personal traits, organizational environment, and educational outcomes. Details are shown in Table 3.

Table 3. Sub-category and Main Findings of TSE Research in higher vocational education

No.	Author(s)	Sub-category	Main findings
1	Agustina et al.,2020	TSE & Teacher's Reflective Practice & Work Engagement & Teacher Trust & Digital Leadership	Teacher self-efficacy plays a key mediating role between principals' digital leadership and teachers' reflective practices. Teacher self-efficacy, in turn, significantly influences their work engagement, which in turn influences their reflective teaching practices.
2	Chou et al.,2010	TCSE & Perceived Organizational & Innovative Climate & Computer Self-Efficacy & Continuous Use of E-Teaching	Teachers' computer self-efficacy mediates the relationship between organizational innovation climate and continued intention to use e-learning. An organizational innovation culture can indirectly promote the sustainability of technology-based teaching by enhancing

			teachers' efficacy.
3	Adebagbo, 2025	TSE & Mastery Experience & Vicarious Experience & Social Persuasion & Physiological and Emotional States	Teacher self-efficacy intrinsically influences teachers' teaching performance through various mechanisms, such as mastery experience, vicarious experience, social persuasion, and emotional state.
4	Gagnon & Dubeau, 2023	TSE & Teachers' professional identity & Training and support mechanisms & Their coping strategies when facing professional challenges	Teacher self-efficacy is closely related to multiple factors, including professional support, teachers' self-perception, and classroom performance.
5	Yang et al., 2024	TSE & Teachers' sustainability knowledge & Teaching readiness	A significant relationship exists between teachers' sustainability knowledge, learning readiness, and self-efficacy.
6	Li, 2025	TSE & Organizational climate	The importance of fostering a favorable organizational climate for enhancing teacher self-efficacy is highlighted, and a significant positive correlation is found between this self-efficacy and teachers' recognition and understanding of VET. Specifically, mediation analysis indicates that teacher self-efficacy moderates the relationship between organizational climate and teachers' perceptions of VET.
7	Mulyono et al., 2023	TSE & Academic writing skills	Teachers who participated in the workshop showed significant improvements in their academic writing self-efficacy.
8	Anwar et al., 2024	TTSE & Teaching effectiveness	There is a positive correlation between teachers' technological self-efficacy and teaching effectiveness, and this increased self-efficacy may promote teaching innovation and student learning outcomes. Furthermore, digital technology use may directly influence teachers' self-efficacy.
9	Ismail et al., 2025	TSE & Satisfaction & Technology acceptance	Increased teacher self-efficacy not only influences educators' use of digital tools but is also directly related to the effectiveness of educational practice and improved student learning outcomes.
10	Lauermann & Berger, 2020	TSE & Teachers' intrinsic motivation & Work engagement & Students' learning motivation and engagement	Teacher self-efficacy plays a significant role in teachers' autonomy-supportive teaching behaviors.
11	Jaedun et al., 2022	TSE & Teaching Effectiveness	The path coefficient between self-efficacy and teaching effectiveness was estimated to be 0.344 and highly significant ($p < 0.001$).
12	Beverborg et al., 2015	TSE & Psychological capital & Teachers' teaching experience & School leadership support	With the support of school leaders, teacher self-efficacy plays a significant mediating role in teachers' participation in professional learning activities.
13	Runhaar et al., 2016	TSE & Career goal orientation & Task interdependence	Teachers' professional self-efficacy is positively correlated with innovative behavior, and this relationship is moderated by career goal orientation and task interdependence.
14	Preechawong et al., 2021	TSE & Career satisfaction & Work experience & Mentoring experience	Mentoring and coaching experiences positively influence teacher self-efficacy and are significantly correlated with job satisfaction.
15	Klaeijssen et al., 2018	Career motivation & Work engagement & Organizational support	Teachers' self-efficacy can mediate the relationship between their work engagement and their innovative behavior.

16	Code et al.,2023	TSE & Career adaptability	Teachers' self-efficacy has a positive impact on their participation in professional learning activities and their ability to cope with teaching changes.
17	Fang & Qi,2023	TSE & School Climate & Teacher Job Satisfaction	Teachers' self-efficacy is significantly positively correlated with job satisfaction. Furthermore, teachers' perceived support from the school and good relationships with colleagues significantly enhance their self-efficacy, which in turn improves their professional identity and work engagement.
18	Siswanto al.,2023	et TSE & leadership & Psychological security climate & Knowledge management & Innovation capability	Self-efficacy has a significant positive impact on teachers' knowledge management and innovative abilities.
19	Kholifah al.,2023	et TSE & Digital Technology Intimacy & Psychological Well-being & Instructional Quality	Self-efficacy has a significant positive impact on digital technology familiarity and psychological well-being, both of which positively predict teaching quality. There is a significant positive correlation between teachers' self-efficacy and teaching quality.
20	Jin et al.,2021	TSE & Professional engagement	Through the implementation of the teacher development program, teachers have significantly improved their self-efficacy in classroom management, self-taught teaching strategies, and student engagement.



Figure 4. Word Cloud Map of Themes Included in the Study

4.2.3.1. Professional practice and engagement

The relationship between teacher self-efficacy and professional practice outcomes was reported in eight studies. Most of these studies employed cross-sectional survey designs ($n = 6$), while two studies adopted intervention or program-based designs. Across these studies, higher levels of teacher self-efficacy were consistently associated with reflective practice, professional engagement, teaching effectiveness, and instructional quality. Several studies further reported mediating or indirect roles of self-efficacy in linking leadership, professional learning, and work engagement. Although findings were

largely consistent, the predominance of cross-sectional evidence limits causal interpretation. Overall, the strength of evidence supporting this theme is therefore moderate.

4.2.3.2. Organizational and contextual factors

Associations between teacher self-efficacy and organizational or contextual variables were reported in six studies, all of which relied on cross-sectional survey designs. These studies highlighted the importance of school climate, organizational support, leadership practices, mentoring experiences, and innovative organizational environments in shaping teachers' self-efficacy beliefs. In several cases, teacher self-efficacy functioned as a mediator or moderator between organizational conditions and outcomes such as job satisfaction, professional identity, and innovative behavior. While the evidence base is coherent, it is methodologically limited by its observational nature. Accordingly, the overall strength of evidence for this theme is assessed as moderate.

4.2.3.3. Psychological and motivational dimensions

A total of five studies examined teacher self-efficacy in relation to psychological and motivational factors, including intrinsic motivation, psychological capital, career adaptability, job satisfaction, and well-being. All studies within this theme used cross-sectional designs, reporting positive and statistically significant relationships between self-efficacy and key psychological resources. These findings suggest that teacher self-efficacy plays an important role in sustaining motivation and psychological resilience among vocational educators. However, given the relatively small number of studies and the absence of longitudinal or experimental evidence, the strength of evidence supporting this theme is considered moderate.

4.2.3.4. Technological competence and digitalization

The integration of technological competence and digitalization into teacher self-efficacy research was addressed in seven studies, including five cross-sectional surveys and two intervention-based or training-focused studies. These studies consistently reported positive relationships between technological or digital self-efficacy and outcomes such as technology adoption, sustained use of e-learning systems, teaching innovation, and instructional quality. Compared with other themes, this area includes a limited but notable presence of intervention evidence, suggesting emerging causal insights. As a result, the overall strength of evidence for this theme is evaluated as moderate to strong.

5. Discussion

This systematic literature review focuses on the current state of research on teacher self-efficacy (TSE) in higher vocational education. Through a systematic analysis of 20 relevant papers, this article explores the definition of TSE, related topics, and key findings. It also examines the publication time and regional distribution of these studies and discusses future research directions in this area. The study found that the number of related studies peaked in 2023. Although it declined between 2023 and 2025, it remained high overall, reflecting the increasing interest in the psychological well-being of vocational education teachers in recent years, in line with global economic development trends. In terms of regional distribution, research is most concentrated in Asia, particularly in Indonesia, followed by China and Malaysia. This phenomenon may be closely related to Presidential Decree No. 68 and the National TVET Reform Strategy, issued by the Indonesian government in 2022. These strategies promote quality improvement in higher vocational education, teacher training, and industry-education integration. For example, the SMK-Center of Excellence Program (SMK-PK) promotes industry-education integration through the teaching factory model (World Economic Forum, 2025). In contrast, research in this area is relatively limited in developed countries with advanced vocational education, such as Germany. Different cultural and national backgrounds lead to significant variations in the function and impact of teachers' self-efficacy beliefs (Bandura, 2002). From a cross-cultural perspective, this regional distribution may also help explain variations in how teacher self-efficacy is conceptualized and reported

across studies. In many Asian contexts, educational systems are often characterized by stronger policy-driven reforms, centralized governance, and a growing emphasis on workforce development. These structural features may place greater expectations on teachers' performance and adaptability, thereby increasing the relevance of self-efficacy as a psychological resource. For example, national initiatives aimed at strengthening technical and vocational education and training (TVET) systems may indirectly promote research on teacher capacity and professional confidence. In contrast, countries with more established vocational education systems, such as those in Europe, may place greater emphasis on institutional stability, apprenticeship traditions, and well-developed training frameworks. In such contexts, teacher self-efficacy may be shaped less by rapid reform pressures and more by professional autonomy and long-standing pedagogical practices. As a result, the function and significance of self-efficacy may differ across cultural and institutional settings. Furthermore, cultural factors such as collectivist versus individualist orientations may also influence how self-efficacy is expressed and measured. In collectivist contexts, teachers' beliefs may be more closely linked to social expectations, institutional norms, and collaborative practices, whereas in more individualistic contexts, self-efficacy may be more strongly associated with personal agency and individual competence. These differences suggest that the interpretation of TSE findings should be situated within specific cultural and educational contexts. Overall, the concentration of studies in Asia not only reflects regional research priorities but also indicates that current findings may be context-dependent. Therefore, future research should adopt more explicit cross-cultural comparative approaches to better understand how cultural, institutional, and policy environments shape teacher self-efficacy in higher vocational education. Therefore, expanding TSE research from a cross-cultural perspective is particularly necessary.

Bandura's (1994) definition of self-efficacy has served as the theoretical foundation for the vast majority of research in this field. Most literature explores self-efficacy within specific teaching contexts, such as technology use (Ismail et al., 2025) and academic writing (Mulyono et al., 2023), focusing on teachers' beliefs and confidence in completing specific teaching tasks and influencing student learning outcomes. Methodologically, research generally relies on the Teacher Self-Efficacy Scale (Tschannen-Moran & Hoy, 2001). However, this scale is a general instrument covering the dimensions of student engagement, classroom management, and teaching strategies, and may not fully capture the characteristics of self-efficacy in specific tasks and fields within higher vocational education (Morris et al., 2017). This limitation is particularly evident in vocational education contexts, where teaching is closely linked to hands-on practice, industry collaboration, and rapidly evolving technical skills. As a result, the use of a generalized instrument such as TSES may overlook task-specific efficacy beliefs that are critical for vocational educators. Furthermore, the continued reliance on TSES across diverse contexts may also contribute to inconsistencies in research findings, as the scale may not adequately reflect contextual variations in teaching environments. This raises questions regarding the construct validity of TSES when applied to specialized educational settings such as higher vocational education. Notably, some studies have begun to adopt more context-appropriate scales, such as the Teacher Artificial Intelligence Self-Efficacy Questionnaire (TASQ), the Computer Self-Efficacy Scale, the Technology Self-Efficacy Scale, and the Academic Creativity and Self-Efficacy Scale (ACSES). These are all based on Bandura's self-efficacy theory and emphasize task-specificity and context-dependence. This shift suggests an emerging trend toward more nuanced and context-sensitive measurement approaches, which may provide more accurate insights into teachers' efficacy beliefs in vocational education settings. Regarding research topics, the literature primarily focuses on TSE and contextual aspects of teaching, such as reflective practice, professional engagement, teaching effectiveness, and teaching quality. It also examines the role of external factors, such as school climate and organizational support, in promoting teacher self-efficacy, as well as psychological factors and motivation. A systematic review shows that teacher self-efficacy has a significant positive impact on improving teaching effectiveness, student engagement, teaching innovation, knowledge management, and psychological well-being, and plays a mediating or moderating role in the relationships between multiple variables. For example, teacher self-efficacy serves as a bridge between principal digital leadership and teacher reflective practice, and between organizational innovation climate and sustained use of e-learning. It also serves as a key link between professional engagement, job satisfaction, and innovative behavior. Furthermore, by enhancing technological self-efficacy, it promotes the adoption of digital tools and improves teaching quality. In summary, this review not only reveals the multi-level influencing paths of teacher self-efficacy, but also

highlights its core position in higher vocational education reform and teacher professional development, and further emphasizes the urgent need for more in-depth cross-cultural and cross-disciplinary research in this field.

6. Implication

In addition to the general implications identified in this review, it is important to translate these findings into more actionable recommendations for higher vocational education institutions. First, institutions should design targeted professional development programs that focus on enhancing teachers' self-efficacy through mastery experiences, such as hands-on teaching practice, industry-based training, and teaching innovation projects. These initiatives can help teachers build confidence in applying both pedagogical and technical skills in real-world contexts. Second, institutional support mechanisms should be strengthened by fostering a positive school climate, including collaborative teaching cultures, mentoring systems, and supportive leadership. Such environments can provide both social persuasion and vicarious learning opportunities, which are essential sources of self-efficacy. Third, vocational institutions should promote the integration of digital tools and emerging technologies into teaching practice, accompanied by structured training and technical support. Enhancing teachers' technological self-efficacy is particularly important in the context of rapidly evolving industry demands. Finally, policymakers and institutional leaders should consider developing context-specific evaluation and support frameworks that reflect the unique characteristics of vocational education, rather than relying solely on general teaching standards. This would allow for more accurate assessment and targeted enhancement of teacher self-efficacy in practice-oriented educational settings.

7. Study limitations

This review also has several limitations. First, we restricted our analysis to empirical studies published in English, excluding books, book chapters, theses, conference papers, and non-English publications. This decision may have introduced language bias and resulted in the omission of relevant studies published in local languages, particularly in regions where vocational and technical education research is frequently disseminated through national journals or institutional outlets. As a consequence, the geographical and cultural coverage of the evidence base may be incomplete.

Second, the exclusion of grey literature, such as policy reports, government documents, institutional evaluations, and doctoral dissertations, may have limited the inclusion of practice-oriented and policy-relevant evidence. Given the applied nature of higher vocational education, important insights into teacher self-efficacy may exist outside peer-reviewed journals and were therefore not captured in this review.

Third, reliance on peer-reviewed journal articles raises the possibility of publication bias, as studies reporting statistically significant or positive findings are more likely to be published than studies with null or negative results. As a result, the strength and consistency of the reported associations between teacher self-efficacy and related outcomes may be overstated.

Fourth, the scope of the review was explicitly focused on teacher self-efficacy in higher vocational education and did not include related psychological constructs such as psychological capital, resilience, or emotional regulation. This focused approach may have constrained the breadth of the theoretical interpretation.

Fifth, this review did not examine demographic or identity-related factors, such as gender, professional rank, disciplinary background, or years of teaching experience, which may moderate the development and effects of teacher self-efficacy. The absence of such analyses limits the depth of contextual interpretation.

Sixth, the use of Google Scholar as a supplementary data source may have introduced potential

sampling bias. Due to the export limitations of Publish or Perish and the non-transparent ranking algorithm of Google Scholar, only the top 1,000 records could be retrieved and screened. As a result, relevant studies that met the inclusion criteria but were ranked beyond this threshold may have been excluded. This truncation may have led to a degree of selection bias. However, this limitation was partially mitigated through comprehensive searches conducted in Scopus and Web of Science, which ensured broader coverage of peer-reviewed literature.

Finally, this study relied on descriptive and thematic synthesis rather than meta-analytic techniques. While this approach is appropriate for mapping research themes and identifying patterns in an emerging field, it does not allow for quantitative estimation of effect sizes or causal inference. Consequently, the findings should be interpreted as indicative of trends and relationships reported in the existing literature, rather than as definitive evidence of effectiveness or causal impact.

8. Conclusion

The primary purpose of this systematic literature review was to explore and synthesize existing research on teacher self-efficacy in the context of higher vocational education. Through a descriptive analysis of 20 studies, this review demonstrates that teacher self-efficacy is a salient psychological construct in vocational education and has received increasing scholarly attention in recent years. The temporal and regional distribution of the literature indicates a growing research focus, particularly in Asian contexts, while evidence from other regions remains relatively limited.

This review addresses an important gap in the literature by foregrounding teacher psychological factors within higher vocational education. The synthesis of findings reveals that teacher self-efficacy in vocational settings is predominantly situated within specific teaching contexts and is closely associated with professional practice, organizational support, and technology integration. In particular, teachers' perceived support from schools and collegial relationships consistently emerge as key factors enhancing self-efficacy, which in turn is linked to stronger professional identity, work engagement, innovative behavior, and teaching quality.

Overall, the findings underscore the importance of strengthening teacher self-efficacy as a strategic lever for improving teaching effectiveness and supporting teachers in adapting to changing instructional models. At the same time, the largely cross-sectional and descriptive nature of the existing evidence highlights important directions for future research. Specifically, future studies should employ more robust designs, such as pre-registered randomized controlled trials or well-designed quasi-experiments, to evaluate the causal impact of targeted professional development interventions on teacher self-efficacy. Longitudinal cohort studies tracking vocational teachers across key career stages would further clarify the developmental trajectories and temporal dynamics of self-efficacy. In addition, cross-cultural research is needed to validate commonly used self-efficacy instruments and establish measurement invariance across regions, thereby improving the comparability and generalizability of findings. Together, these directions can strengthen the empirical foundation of research on teacher self-efficacy and better inform evidence-based policy and practice in vocational education.

Data Availability Statement

The data supporting the findings of this study, including the CCAT quality appraisal file, the literature extraction table for the 20 included studies, and the full-text screening log with reasons for exclusion, are openly available in Zenodo at <https://doi.org/10.5281/zenodo.18288355>

Declaration of AI Use

Artificial intelligence-based tools were used during the preparation of this manuscript to assist with language polishing and structural refinement. All academic judgments, data interpretation, theoretical

framing, and final content decisions were made by the authors. The authors take full responsibility for the integrity and originality of the manuscript.

References

- Adebagbo, A. O. (2025). Artificial Intelligence Integration in Teaching and Learning: Investigating Retirement-date and Self-efficacy of in-service Teachers in Higher Education Institutions. *Rima*, 4(1).
- Agustina, R., Kamdi, W., Hadi, S., & Nurhadi, D. (2020). Influence of the principal's digital leadership on the reflective practices of vocational teachers mediated by trust, self efficacy, and work engagement. *International Journal of Learning, Teaching and Educational Research*, 19(11), 24-40.
- Ampofo, J., Bentum-Micah, G., Xusheng, Q., Sun, B., & Mensah Asumang, R. (2025). Exploring the role of teacher empathy in student mental health outcomes: a comparative SEM approach to understanding the complexities of emotional support in educational settings. *Frontiers in Psychology*, 16, 1503258.
- Anwar, C., Sofyan, H., Ratnaningsih, N., & Am, M. A. (2024). Digital Technology Practices for Vocational Teachers in the Industrial Revolution 4.0: Mediating Technology Self-Efficacy. *Journal of Pedagogical Research*, 8(1), 172-190.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.
- Bandura, A. (2002). Social cognitive theory in cultural context. *Applied psychology*, 51(2), 269-290.
- Bandura, A., & Wessels, S. (1994). *Self-efficacy* (Vol. 4, pp. 71-81). na.
- Chou, C. M., Hsiao, H. C., Shen, C. H., & Chen, S. C. (2010). Analysis of Factors in Technological and Vocational School Teachers' Perceived Organizational Innovative Climate and Continuous Use of E-Teaching: Using Computer Self-Efficacy as an Intervening Variable. *Turkish Online Journal of Educational Technology-TOJET*, 9(4), 35-48.
- Chou, C. M., Hsiao, H. C., Shen, C. H., & Chen, S. C. (2010). Analysis of Factors in Technological and Vocational School Teachers' Perceived Organizational Innovative Climate and Continuous Use of E-Teaching: Using Computer Self-Efficacy as an Intervening Variable. *Turkish Online Journal of Educational Technology-TOJET*, 9(4), 35-48.
- Code, J., Moylan, R., Forde, K., & Ralph, R. (2023). Teachers' Sense of Efficacy During a Time of Crisis. *Canadian Journal of Science, Mathematics and Technology Education*, 23(3), 538-558.
- Fang, J., & Qi, Z. (2023). The influence of school climate on teachers' job satisfaction: The mediating role of teachers' self-efficacy. *Plos one*, 18(10), e0287555.
- Fackler, S., & Malmberg, L.-E. (2016). Teachers' self-efficacy review. *Educational Psychology Review*, 28, 769–791.
- Gagnon, N., & Dubeau, A. (2023). Building and maintaining self-efficacy beliefs: A study of entry-level vocational education and training teachers. *Vocations and Learning*, 16(3), 511-532.
- Gao, X., Yu, C., Li, H., & Tao, J. (2023). The current situation and intervention measures of mental health literacy of vocational college early childhood education students from the perspective of preschool integrated education—a disease perception perspective based on attention deficit hyperactivity disorder in children. *Applied & Educational Psychology*, 4(11). <https://doi.org/10.23977/appep.2023.041112>

- Gunawardena, H., Leontini, R., Nair, S., Cross, S., & Hickie, I. (2024). Teachers as first responders: classroom experiences and mental health training needs of Australian schoolteachers. *BMC public health*, 24(1), 268.
- Hilman, H., Wahyudin, D., & Rusman, R. (2020). A study of multi entry and multi exit education system in increasing vocational high school graduates skills. *Applied Science and Innovative Research*, 4(2), p70. <https://doi.org/10.22158/asir.v4n2p70>
- Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of educational psychology*, 105(3), 774. <https://doi.org/10.1037/a0032198>
- Ismail, M. H., Ali, S. H. S., Husin, H. S., Hassan, S., & Ali, S. N. D. M. (2025). From Adoption to Continuance: Understanding Self-Efficacy's Role in TVET Digital Learning. *Journal of Technical Education and Training*, 17(2), 128-136.
- Jaedun, A., Nurtanto, M., Mutohhari, F., Majid, N. W. A., & Kurdi, M. S. (2022). Mediating Role of Teachers' Self-Efficacy and Psychological Capital in Determining Success during Learning Transition Periods in higher vocational education. *Journal of Education and E-Learning Research*, 9(3), 207-215.
- Jin, X., Tigelaar, D., van der Want, A., & Admiraal, W. (2023). The effects of a teacher development programme in chinese vocational education on the efficacy and professional engagement of novice teachers. *Journal of Education for Teaching*, 49(2), 252-265.
- Kamel, A., Haridi, H. K., Alblowi, T. M., Albasher, A. S., & Alnazhah, N. A. (2020). Beliefs about students' mental health issues among teachers at elementary and high schools, Hail Governorate, Saudi Arabia. *Middle East current psychiatry*, 27(1), 30.
- Kholifah, N., Kurdi, M. S., Nurtanto, M., Mutohhari, F., Fawaid, M., & Subramaniam, T. S. (2023). The role of teacher self-efficacy on the instructional quality in 21st century: A study on vocational teachers, Indonesia. *Int J Eval & Res Educ ISSN*, 2252(8822), 8822.
- Kholifah, N., Kurdi, M. S., Nurtanto, M., Mutohhari, F., Fawaid, M., & Subramaniam, T. S. (2023). The role of teacher self-efficacy on the instructional quality in 21st century: A study on vocational teachers, Indonesia. *Int J Eval & Res Educ ISSN*, 2252(8822), 8822.
- Klaeijssen, A., Vermeulen, M., & Martens, R. (2018). Teachers' innovative behaviour: The importance of basic psychological need satisfaction, intrinsic motivation, and occupational self-efficacy. *Scandinavian Journal of Educational Research*, 62(5), 769-782.
- Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy. *Educational Psychology Review*, 26(3), 419-452.
- Klassen, R. M., et al. (2009). Teacher efficacy cross-cultural. *Journal of Educational Psychology*, 101(3), 673-689.
- Lauermann, F., & Berger, J. L. (2021). Linking teacher self-efficacy and responsibility with teachers' self-reported and student-reported motivating styles and student engagement. *Learning and Instruction*, 76, 101441.

- Li, Y., Li, T., Liu, J., Hu, Z., & Qi, Z. (2025). Decoding vocational teachers' professional identity: Organizational climate and teacher self-efficacy as key factors in the Chinese educational context. *International Journal of Educational Research*, 132, 102651.
- Lucas, B., et al. (2012). *How to teach vocational education*.
- Moher, D., et al. (2009). PRISMA statement. *PLoS Medicine*.
- Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational psychology review*, 29(4), 795-833.
- Mulyono, H., Ningsih, S. K., Fausia, F., Setiawan, H., Ibarra, F. P., & Mukminin, A. (2023). Developing an academic writing creativity and self-efficacy among Indonesian TVET instructors: Evaluating an online genre analysis-based academic writing workshop. *Cogent Education*, 10(2), 2237319.
- Na, C., & Isa, Z. M. (2024). Exploring the influence of teacher self-efficacy on teaching quality in higher vocational education. *Journal of Digitainability, Realism & Mastery (DREAM)*, 3(07), 16-27.
- OECD (2021). *Teachers and School Leaders as Lifelong Learners*.
- O'Toole, C. (2023). The Role of Teachers in Supporting Students' Mental Health. *Teachers and Curriculum*, 23(1), 89-99.
- Oude Groote Beverborg, A., Slegers, P. J., & Van Veen, K. (2015). Promoting VET teachers' individual and social learning activities: the empowering and purposeful role of transformational leadership, interdependence, and self-efficacy. *Empirical research in higher vocational education and training*, 7(1), 5.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372.
- Perera, H. N., et al. (2019). Self-efficacy meta-analysis. *Educational Psychology Review*, 31, 1–45.
- Preechawong, S., Anmanatrakul, A., Pinit, P., Koul, R., & Easter, M. A. (2024). Relationship between mentoring and coaching experience, teaching self-efficacy and job satisfaction of vocational school teachers in Thailand. *Educational Studies*, 50(5), 722-742.
- Reinke, W. M., Herman, K. C., Stormont, M., Yang, W., & Wan, C. (2025). Supporting children's mental health in school over a decade later: Current teacher perspectives. *School Psychology*.
- Runhaar, P., Bednall, T., Sanders, K., & Yang, H. (2016). Promoting VET teachers' innovative behaviour: exploring the roles of task interdependence, learning goal orientation and occupational self-efficacy. *Journal of Vocational Education & Training*, 68(4), 436-452.
- Sisask, M., Värnik, P., Värnik, A., Apter, A., Balazs, J., Balint, M., ... & Wasserman, D. (2014). Teacher satisfaction with school and psychological well-being affects their readiness to help children with mental health problems. *Health education journal*, 73(4), 382-393.

Siswanto, E., Samsudi, E. S., & Sutopo, Y. The role of psychological security climate, leadership, and self-efficacy on teachers capability through knowledge management. *Int J Eval & Res Educ ISSN*, 2252(8822), 8822.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, 17(7), 783-805.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of educational research*, 68(2), 202-248.

UNESCO (2022). *Transforming TVET for sustainable development*.

Victor, L. (2008). Systematic reviewing. *Social research update*, 54(1), 1-4.

Vieluf, S., et al. (2013). Teaching practices and self-efficacy. *OECD*.

Wang, P. and Cao, Z. (2024). Exploring the characteristics and path of internationalization development of vocational education based on the information diffusion model. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-0379>

Whitley, J., Smith, J. D., & Vaillancourt, T. (2013). Promoting mental health literacy among educators: Critical in school-based prevention and intervention. *Canadian Journal of School Psychology*, 28(1), 56-70.

Wheelahan, L. (2015). Vocational education theory.

World Economic Forum. (2025, May 15). Indonesia's green jobs and vocational education transformation. World Economic Forum. <https://www.weforum.org/stories/2025/05/indonesia-green-jobs-vocational-education-training/>

Yang, W., Chinedu, C. C., Chen, W., Saleem, A., Ogunniran, M. O., Ñacato Estrella, D. R., & Vaca Barahona, B. (2024). Building capacity for sustainability education: An analysis of vocational teachers' knowledge, readiness, and self-efficacy. *Sustainability*, 16(9), 3535.

Yang, W., Chinedu, C. C., Chen, W., Saleem, A., Ogunniran, M. O., Ñacato Estrella, D. R., & Vaca Barahona, B. (2024). Building capacity for sustainability education: An analysis of vocational teachers' knowledge, readiness, and self-efficacy. *Sustainability*, 16(9), 3535.

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy. *Review of Educational Research*, 86(4), 981-1015.

Zhang, Y. (2017, June). The present situation and Countermeasures of teachers' psychological health in Higher Vocational Colleges. In *2017 2nd International Conference on Education, Sports, Arts and Management Engineering (ICESAME 2017)* (pp. 482-486). Atlantis Press.

Stakeholder Perceptions of EAP Curricula in Higher Education: A Systematic Literature Review

ABSTRACT

Authors: Abuova Janar Salimovna & Sharifah Intan Sharina Syed Abdullah*
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:
sharifahintansharina@upm.edu.my

Received: 31/12/2025 26 August 2024 |

First Revision: 14/1/2026 11 February 2025 |

Second Revision: 15/4/2026: 28 May 2025 |

Accepted: 15/5/2026 30 June 2025

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

A systematic literature review (SLR) was conducted to study stakeholders' perceptions regarding the English for Academic Purposes curriculum. The SLR focuses on the insights regarding curriculum development, such as relevance, challenges and effectiveness of it from the experiences of students, instructors (both language and content) and administrators. As EAP plays a central role in developing students' academic communication skills and supporting academic success in multilingual contexts, understanding stakeholder insights is essential for informed curriculum development. SLR covered empirical, peer-reviewed studies published between 2000 and 2025. Four searching engines were used: Scopus, ERIC, ScienceDirect, and SpringerLink. Study selection followed the PRISMA framework, and the inclusion criteria were based on the PICO model. Data were extracted into a standardized table and thematically analysed following Braun and Clarke's (2006) six-phase framework, with quantitative findings summarized descriptively. Findings indicate that stakeholders generally recognize the value of EAP curricula in promoting academic literacy and communicative competence. However, several gaps persist. Students often report limited transferability of EAP skills to discipline-specific academic contexts. Instructors highlight challenges related to curriculum flexibility, pedagogical alignment, and institutional support, while administrators emphasize the need for ongoing evaluation and collaboration to maintain curriculum relevance. The review concludes that stakeholder engagement is critical in ensuring that EAP curricula effectively address diverse academic and disciplinary needs. It recommends participatory curriculum review processes, stronger alignment between EAP content and academic tasks, and institutional policies that foster collaboration among all stakeholders. Additionally, there is emerging theme for further studies incorporating sustainability competencies integration.

Keywords: English for Academic Purposes (EAP), stakeholder perceptions, higher education, curriculum development, systematic review, thematic analysis.

1. Introduction

English for Academic Purposes (EAP) is one of the central parts of the English medium of instruction (EMI) teaching in Higher Educational Institutions (HEIs). The main aim of the module is to equip learners with essential skills supporting their academic success in an English-medium context. EAP equips students at HEIs with important skills such as communicative competence, critical thinking, reading academic texts, writing research-oriented writing, and presenting and participating in seminars (Hyland, 2006; Hu & Lei, 2014). The rise of internationalization and the adaptation of the EMI at universities increases demand for well-structured EAP programs (Ding & Bruce, 2017). EAP is not only about teaching terminology of the specific field or in a general context, but it is also more about the discourse practices of the academic world. It supports students to understand and participate in the chosen discipline (Hyland, 2016). As a result, EAP is a bridge that helps students to have smooth academic acculturation and navigate them in their disciplinary expectations (Floyd, 2015). Additionally, it serves for international and second language students as cultural and linguistic scaffolding, providing an equitable transition into academic life (Walqui, 2006). Therefore, when students transfer from the comfortable environment of their own language into English, the support of EAP both linguistically and disciplinary makes it inevitable.

Even though the importance of EAP is mentioned a lot in the scholarly world, the implementation of it in the curriculum of higher educational institutions all the time differs across regions and even institutions. Such factors as institutional goal, resources, teacher's expertise, and administrators' competence influence the designed program (Aksit & Aksit, 2022). This divergent attitude toward the program shows the importance of the evaluation of how EAP curricula are perceived and experienced by stakeholders.

Understanding stakeholders' perceptions, especially students, instructors, and administrators, will ensure that the program is relevant and effective. The needs of students will provide insight regarding how well the program is considering their academic, linguistic needs, and instructors' perspectives present pedagogical and practical challenges in the course delivery (Dou, 2024). On the other hand, administrators present administrators' perspectives on the curriculum goals, policy priorities, and resource division (Iranmehr & et. al, 2018). The combination of all those perspectives shapes the high quality of EAP teaching.

However, current studies on stakeholders' perceptions are fragmented. A lot of studies are focusing on one group of stakeholders (mostly students) or on local case studies, and there is almost no integration of the findings across the various contexts (Bakogiannis & Papavasiliou, 2025). Although a number of earlier systematic literature reviews have discussed EAP and academic literacy, most of them have addressed the issue through a prism of pedagogical practices (Csiernik & et.al, 2021). The wider literature on EMI has been charted by other researchers without a specific consideration of stakeholder views of EAP curricula (Pomat, 2025). This current review builds upon this previous study in three aspects; first, it adopts a multi-stakeholder approach, which, at the same time, incorporates the views of students, instructors, and administrators; second, it synthesizes the literature published to date 2000-2025, thus updating the previous reviews of EAP curriculum syntheses. As a result, this research aims to review collective insights of multiple stakeholders to understand how EAP curricula are conceptualized, valued, and critiqued in different higher educational contexts. Specifically, it addresses the following research questions:

1. How do different stakeholders (students, instructors and administration) perceive the purpose, design, and effectiveness of EAP curricula in higher education?
2. What similar themes, challenges, and contextual differences emerge from these perceptions?
3. What consensus recommendations for EAP curriculum improvement emerge across stakeholder groups and institutional contexts?

The systematic analysis of the stakeholders' viewpoints across various contexts will provide a holistic understanding of EAP curriculum design and evaluation. The review highlights the importance of

critical thinking, problem-solving skills and research writing for the improvement of education and the role of the EAP in their integration from the first steps of higher education by integrating them in the course curricula.

1.1. Theoretical and Conceptual Framework.

This section is a theoretical and conceptual framework of the current systematic review. The framework has three functions firstly; to operationalize the key constructs; secondly, to chart the relationships between the constructs based on the existing curriculum theory; and third, to describe the way the framework informs the interpretation of the perceptions of stakeholders throughout the literature reviewed.

This review is guided by two key constructs, including stakeholder perception and EAP curriculum. The operational definition of stakeholder perception can be described as cognitive and affective judgment that individuals have about the purpose, design, delivery, and results of an EAP curriculum. Based on the literature in educational evaluation (Al Issaei, 2018; Kirkpatrick, 1998), this definition covers three dimensions, which are related to each other such as cognitive, affective and evaluative. Notably, this review acknowledges that not all stakeholders have the same perceptions. Each of the students, instructors, and administrators is in different positionalities in the educational system possessing varied interests, duties and access to information. Thus, they can be similar or dissimilar in perceptions meaningfully and the patterns of similarity and dissimilarity are considered data of analytical importance and not error of measurement.

The operationally defined EAP curriculum is the planned and implemented educational experience that aims at providing students with the academic language and literacy skills to succeed in English-speaking higher education. In order to systematise the operationalisation of this construct, the review uses the framework of curriculum development proposed by Richards (2001) as the main theoretical framework. Richards rationalises curriculum as consisting of a number of interacting phases that constitute a dynamic and reiterative system. These stages involve needs analysis, setting of goals and objectives, selection and sequencing of contents, instructional delivery and methodology, materials development, assessment design and programme evaluation. In this context a curriculum is not considered a fixed document but it is a living process that develops through the continuous processes of planning, implementation and refining. The framework also stresses that an effective curriculum development should focus on the intended curriculum (what is planned) and the enacted curriculum (what actually occurs in the classrooms) and the interrelationship between the two.

Mapping Relationships between Constructs with the Framework of Richards.

The curriculum development framework created by Richards (2001) can help in gaining a systematic view of the relationship between the opinions of the stakeholders and the quality of EAP curriculum. In this context, the stakeholder perceptions are perceived to work at various levels through the curriculum cycle.

To begin with, stakeholder perceptions shape the needs analysis phase that is described by Richards as the initial phase of curriculum development. Perceptions of students about their own linguistic and academic needs, perceptions of instructors about classroom realities and student challenges and perceptions of administrators about institutional priorities and resource constraints are all critical input to the process of deciding what an EAP curriculum should cover. As these perceptions are assembled systematically and put together, then the resultant curriculum will be potentially more relevant and responsive. They may be neglected or the groups of stakeholders may have different perceptions that lack ways of reconciling hence the curriculum may end up being out of sync with the real needs.

Second, curriculum delivery and curriculum methodology are shaped and have an influence on stakeholder perceptions. Richards underlines that the curriculum implemented by teachers and students in classrooms, the curriculum actually being implemented, can be quite different as compared to the

planned curriculum. The perceptions of instructors about the feasibility, student motivation, and available resources affect the way they understand and adjust curriculum plans. Equally, perceptions of students with regard to relevance, difficulty, and engagement to the task influence their engagement and the learning outcomes. These perceptions form feedback loops: positive perceptions are likely to support the continued effective delivery, whereas negative perceptions can result in disengagement, resistance, or demands to change the curriculum.

Third, the stakeholder perceptions can be considered as informal curriculum evaluation. The programme evaluation is an important last step that is included in the framework by Richards, but formal evaluation mechanisms are usually limited or not very common. Without them, the stakeholder perceptions would act as the daily, continuous measures of the quality of the curriculum. The high degree of congruence between student, instructor and administrator perceptions (such as all three groups agreeing that the EAP curriculum prepares students effectively to write about disciplinary subjects) is indicative of curriculum coherence and effectiveness. On the other hand, systematic discrepancies, i.e. when administrators think that the curriculum works and students and instructors tell them that there are serious issues in the curriculum, indicate that there may be communication, implementation, or resource-distribution failures in the system.

The present systematic review should be viewed as an attempt to explain the rationale of the interpretation through the adoption of the Richards (2001) framework of curriculum development. Since the review is a compilation of results of the studies that present the perceptions of the stakeholders about EAP curricula, the framework can provide a set of conceptual anchors that can be used to structure and interpret different results in different institutional settings.

Where a study reports that students consider EAP writing tasks irrelevant to their disciplinary tasks, the framework interprets this result as a perception found at both the needs analysis (were student needs correctly assessed) and curriculum delivery (is the implemented curriculum appropriate to disciplinary realities) points. Where a paper indicates that instructors believe that they do not have sufficient training to teach discipline-specific academic writing, the framework would understand this as a perception about the resourcing and implementation stages of curriculum development, indicating that there is a mismatch between curriculum plans and instructor capacity. When an experiment results in administrators putting standardised testing first and students and instructors demand more authentic assessment, the framework views this difference as a lack of coherence between stakeholder groups that points to possible issues in the curriculum evaluation and revision processes.

In this interpretive process, the framework created by Richards is more of an orienting device and not a strict template. It enables the review to systematically classify the perceptions of the stakeholders based on the stage of curriculum development that they majorly relate to-needs analysis, goal setting, content selection, delivery, materials, assessment or evaluation. Meanwhile, the framework is also open to emerging themes not initially foreseen by the theoretical framework, including the incorporation of sustainability competencies into EAP curricula. Once these emergent themes appear inductively through the literature reviewed, they are presented as exploratory results potentially extending or challenging the current framework, thus indicating future directions of theoretical development.

2. Methodology

2.1. Review Design

The research combines a systematic review design with narrative thematic synthesis in the research to explore how university stakeholders perceive the EAP curriculum in higher education. The systematic literature review was selected to guarantee methodological, replicability, and transparency during the selection and analysis process. As stakeholders' perceptions are complicated and depend on the context where they happen, it was decided to study them qualitatively (Fendt, 2025). Narrative synthesis will

help to analyze different data sources, including qualitative interviews, surveys, and mixed-method studies in the diverse EAP curriculum research.

The review covers articles published over a 25-year period, starting from 2000 to September 2025. This period covers the most important milestones in the EAP development, such as the expansion of the EMI programs, especially in post-soviet countries. Additionally, it coincides with the evolution of the curriculum standards, assessment, institutional expectations, and HEIs' advancements over the years. The restriction of the review will cover the last two and a half decades, and cover the most modern trends without eliminating crucial shifts in the stakeholders' needs and curriculum design frameworks over the years.

2.2. Study Selection Process

The research paper selection process went through a multi-stage screening and selection procedure. The process followed PRISMA 2020 guidelines to emphasize accuracy, transparency, and replicability in the systematic review of the material (Page et al., 2021). The stage of identification was across four major databases, such as Scopus, ERIC, ScienceDirect, and SpringerLink. All those databases are easy to access, have full possibilities to gain the text, and there is a high chance of selecting a reliable and valid research paper that will avoid predatory journals. The first stage of the selection procedure started from the manual selection through the title and abstracts of the journals, resulting in 318 records from ERIC (24 retained for screening), 227 from Scopus (33 retained), 204 from SpringerLink (9 retained), and 796 from ScienceDirect (28 retained). The second stage of the selection included reading the whole article, resulting in 24 records from ERIC (6 retained for qualitative study), 33 from Scopus (6 retained for qualitative study), 9 from SpringerLink (0 retained), and 28 from ScienceDirect (9 retained for qualitative study). Resulting in 21 papers at the final stage. During this period, all the duplicated articles or irrelevant records were excluded, resulting in overall 21 extractions.

The main aim of the title screening stage was to remove studies unrelated to the scope of EAP, such as general English instruction, K-12 English, training programs not connected to an academic context, and others. Only titles connected to EAP, curriculum development, or stakeholders' perceptions in higher education were considered. The next stage of abstract screening was conducted with the help of the designed criteria to guarantee that the remaining studies include the involvement of university-level stakeholders, consider perceptions and experiences, and are relevant to the EAP curriculum. The paper was not included if the abstract showed no information about perceptions and was limited to the language proficiency outcome. Also, if the population was not relevant to the study, it was not included in the list.

Table 1. Prisma Diagram

Stage	Database	N of identified articles	N of eliminated articles	N of duplicated articles	N of articles included for further analysis
Identification and screening for title and abstract	ERIC	318	296	not checked for duplication	24
	Scopus	227	194		33
	SpringerLink	204	195		9
	ScienceDirect	796	768		28
Screening for the whole text	ERIC	24	18	3	6
	Scopus	33	27		6
	SpringerLink	9	9		0

	ScienceDirect	28	19		9
Reports assessed for quality	ERIC	6	0		6
	Scopus	6	0		6
	ScienceDirect	9	0		9
Included studies	21				

The next stage followed was full-text screening, where each article was read in detail. The exclusion at this stage happened when there was no voice of the stakeholders, when the article did not contain empirical data, or when participants were not part of a HEI. Following the PRISMA model, all the decisions were documented systematically to ensure that the transparency of the study is maintained. After the selection process, each selected article was checked for quality appraisal. It was done to assess the clarity of research design, transparency in sampling, adequacy of data collection, richness of contextual description, usage of trustworthiness strategies, implementation of ethical considerations, and overall relevance to the research review questions. After the quality appraisal, all 21 selected articles were kept for further analysis. The assessment was provided with the rubric and assessed as yes, no, or unclear options, providing a chance for numerical scoring.

In Table 1, it is possible to see an illustrated process in the PRISMA model of the systematic literature review selection procedure, which included screened, assessed, and included articles. This systematic process ensured that only studies directly connected to the aims of the review were selected for the final stage.

2.3. Search Strategy

The searching strategy of the study was designed based on the step-by-step approach to ensure comprehensive and detailed coverage of research related to the stakeholders' perception regarding the English for Academic Purposes. There were selected 4 databases due to their strong coverage of the social science and applied linguistics: Scopus is a database which is famous for its reliability and validity in the selection process of the quality journals especially those having strong connection with social science and applied linguistics; ERIC, specialized education oriented database focused on language learning and pedagogy; Science Direct, one of the broadest databases which contains the most research papers about academic and professional education context; and the Last one is SpringerLink, which contains peer-reviewed and open access journals connected with language teaching, higher education and curriculum development. One more criterion that was considered to maintain scholarly rigour was limiting the journal selection only with peer-reviewed, English language journal articles.

To maximize the sensitivity and specificity of the searching terms, it was decided to use Boolean operators. The initial searching key word line used across databases was: ("Perceptions" OR "Perspectives" OR "Attitudes") AND ("Stakeholders" OR "Teachers" OR "Students" OR "Administrators") AND ("EAP" OR "English for Academic Purposes") AND ("Curriculum" OR "Syllabus" OR "Program" OR "Course") AND ("Higher Education" OR "University"). If the searching platform had no Boolean capabilities, there was a simplified version of the key terms such as "Perceptions", "Stakeholders", "EAP", and "Curriculum". All the filter settings were set to restrict the search to peer-reviewed publications written in English in the period between 2000 and 2025. The exact searching string with applied strings is available in Appendix I.

After retrieving records from the platform during the second stage of screening, duplicate entries were removed manually. After checking the second eligibility criteria and quality checking stages, the first pool of articles was generated for full-text assessment.

After retrieving records from all databases, the results were merged, and duplicate entries were removed using a reference management system. The remaining studies were then screened for relevance through title and abstract review according to the predefined eligibility criteria. This process generated the initial pool of articles for full-text assessment.

2.4. Eligibility Criteria

The adaptive version of the PICO framework was developed for eligibility criteria. It was tailored for qualitative, quantitative and mixed-methods research in education. The first population component focused on university stakeholders who are directly involved in EAP curriculum design, decision-making, teaching and learning. The population included students, instructors, curriculum developers, and administrators. Studies which included information regarding K-12, general English or non-academic English subjects were excluded. The second component, Intervention, was selecting stakeholders' perceptions, experiences, attitudes or evaluations related to articles. There were studies about curriculum design, implementation, delivery, and alignment of needs and curriculum outcomes. Studies focusing only on language proficiency or other unrelated pedagogical elements were not considered. The next element of comparison supports the inclusion of studies offering comparative perspectives across stakeholder groups, programs and institutional context, provided that all the comparisons are directly related to EAP. The last outcome domain was the same way focusing on studies reporting stakeholders' perceptions, curriculum effectiveness, relevance and recommendation. The articles without or with limited orientation on perception were excluded. The final and most important eligibility criterion was only in English, peer-reviewed journal articles, which were published between 2000 and August 2025. All empirical studies, including qualitative, quantitative and mixed-methods, were eligible as long as they met the provided criteria.

The screening process took place in two phases because of the application of these criteria. The title and abstract screening were done in the first stage. At this point, 1453 records were filtered out (Eric 296, Scopus 194, SpringerLink 195, ScinceDirect 768). The most common exclusion criteria were: (a) the research included studies about K-12 or non-university setting, (b) the research included studies about general English or English for Specific Purposes (ESP) without explicit focus on the outcomes in language proficiency, and (c) the research included studies that reported stakeholder perceptions. In the second stage, full-text screening was performed on the remaining 95 articles. Out of the total number of articles screened 73 were full-text articles that were excluded (Eric 18, Scopus 27, SpringerLink 9, ScinceDirect 19).. The rationale behind this exclusion was as follows (counts): no explicit stakeholder perception data were given, non-EAP language teaching settings were considered, theoretical or opinion articles lacked empirical data, and did not have sufficient relevance to curriculum design, delivery or assessment (3). This two-step screening procedure ensured that all the studies that were selected directly answered the main research question of the review: How do stakeholders perceive the EAP curriculum in higher education? The full access to the eligibility criteria dataset could be visible through this <https://doi.org/10.5281/zenodo.18489466> link.

2.5. Quality Assessment

After the final selection procedure, a structured quality assessment of the articles was implemented to evaluate the methodological rigour and relevance of the studies chosen for the review. The purpose of the process was not to exclude studies. Mainly, it was aimed to enhance transparency in data interpretation and prove the reliability of the evidence provided. In order to make this process in a very smooth way there was a quality appraisal rubric. The rubric was adapted from the qualitative checklist of the Joanna Briggs Institute critical appraisal tools (Joanna Briggs Institute, 2017). The focus of the review was on studies exploring stakeholders' perceptions of EAP curricula.

The rubric contained several important elements. The first domain was Study Design appropriateness, which assessed the clarity of the selected papers' research design along with its connection with the research question. The second and third domain is regarding sampling and sampling method, which identified stakeholder groups, adequacy of the sample size, transparency of sampling and recruitment

of the procedure. Data Collection is the fourth domain evaluating the transparency of the data collection tool, including ethical issues like confidentiality. The next domain, Data Analysis Quality, checked that the article has a clear and systematic data analysis process, including evidence of credibility such as triangulation and validation. The final section is about contextual theoretical evidence, where it was assessed that context was contextualized enough, and the chosen theoretical framework was clearly connected to the study.

Each criterion was assessed using a three-point scale: 2 (YES), 1 (PARTIAL) and 0 (NO) i.e. fully, partially, or not addressed. A cumulative score was computed on each study included based on all criteria. The highest possible score was 12 and the lowest possible score was 0. The application of the rubric is critical because it demonstrates the diversity of methodology of the studies and their difference in the quality of reporting (Azarian et al., 2023). In total, 21 articles were evaluated in terms of their quality. The minimum score of the chosen articles was 6 and the maximum was 12. To explain the score allocation: 13 studies had a score of 10-12 (high quality), 3 studies had a score of 7-9 (moderate quality), 5 studies had a score of 6 (the minimum threshold). It was originally a part of the quality assessment to inform the interpretation of findings but was not meant to be an exclusion criterion. Nevertheless, in order to be methodologically relevant, it was agreed that the study with a score lower than 50% (that is, less than 6 out of 12) would not be included. This criterion was clear cut: those studies with a score of 5 or less were eliminated; those studies with a score of 6 or higher were retained.

After this threshold was applied, no studies were eliminated based on quality criteria as all 21 studies initially identified met this criterion with 6 points or higher. The rationale behind the inclusion of all studies was that, despite the limited methodological capabilities of such research, it may still yield valuable information, especially when the aim is to chart the diversity of stakeholder perceptions in different EAP contexts. No formal kappa statistic on the inter-rater reliability on the quality appraisal was carried out because the appraisal was carried out by one reviewer who was then audited (as outlined in the coding reliability section). Appendix II presents the list of 21 selected articles, including the quality rubric scores of each article.

2.6. Data Extraction

The data extraction process was focused on 5 analytical elements, which included the most important parts of the study, which would help to respond to the key research question of SLR. The dimensions are Participants, EAP Focus or Intervention, Key Findings, Positive/Negative Perceptions, and Recommendations. Keeping fixed data extraction provided consistency in the data interpretation in the further data analysis process. Moving to the first section of the extracted data, it is important to note that the participants contained a variety of stakeholder representatives, including undergraduate and postgraduate students, EAP and Content instructors, curriculum developers, and institution administrators. This wide variety of stakeholders give an opportunity to see broader perceptions and experiences regarding EAP provision in the higher educational context. The section about EAP focus and intervention included data about curriculum design, course delivery, assessment, skill integration, teacher preparation for the EAP teaching role and academic literacy. There were also some studies which were focusing on institutional reforms, policy changes and the transition of the Language teaching modules. The next section regarding Key findings extracted data about curriculum relevance, its connection with academic expectations, readiness of the instructors, lesson materials and the way it was assessed. The Positive and negative perspectives section observed and documented information that was regarding positive trends, such as improved skills, increased awareness of academic notions, while negative perceptions included such data as outdated materials, heavy workload limiting teachers' professional development and others. Finally, the last section focused on various suggestions which were recognized in the articles and highlighted the improvement of EAP. Organizing data within these five dimensions ensured a structural and transparent syllabus synthesis that captured different perspectives of the stakeholders. The detailed data extraction field could be found in the [10.5281/zenodo.18513532](https://zenodo.org/record/105281).

2.7. Data Analysis and Synthesis

The process of analysis was systematic and rigorous to ensure that no important data from the reviewed literature would be missed or wrongly interpreted. The analytical approach, which was used for thematic analysis, was guided by the framework developed by Braun and Clark (2006). This model contains six phases: data familiarization, initial coding, theme generation, theme review, definition of themes and final reporting of the findings. This structure is flexible but at the same time robust and will clearly state stakeholders' perception regarding EAP program development. The usage of this approach provided an inductive and theoretically supported analysis, which was directly retrieved from the extracted data. The detailed extraction table and the coded materials could be found in [10.5281/zenodo.18513532](https://zenodo.org/record/18513532).

The thematic coding was performed by a single coder who used the thematic codebook over the whole dataset of 21 studies included. To maintain quality control and to determine the validity of the coding procedure, an independent audit by a second researcher who was not involved in the initial coding was done. The second researcher examined a purposively sampled coded data of 6 out of 21 studies. Sampling was selected to reflect the various stakeholder views (students, instructors, administrators) and a geographical distribution (East Asia, Western countries, and the small number of available studies in other areas).

The second researcher did the audit process to ensure every given code was relevant and eligible according to the operational definitions and indicators of the codebook. In particular, regarding each coded segment, the second researcher determined: (a) whether the code was correct in terms of the content of the extracted finding, and (b) whether the code was in line with the thematic category as specified in the codebook. The formal kappa statistic was not computed, because the audit was done on content relevance and eligibility as opposed to chance-corrected agreement; in this case, percent agreement is a clear and sufficient measure of coding consistency due to the design of the audit and the process of consensus.

Any inconsistencies and unclear situations found in the course of the audit were resolved by exchanging views in the online meetings between two researchers. During these consensus discussions, every dispute was discussed, the pertinent text re-read, and the codebook definitions were consulted in order to come to a common decision. No third reviewer was needed, because all the disagreements were overcome by means of the dialogue. This method of a single primary coder, then an independent audit, and consensus based verification offered a sound system of assuring consistency of coding and reduced individual bias. The coding consensus formed the basis of all future thematic synthesis that was reported on in this review.

The thematic synthesis was guided by six-phase framework developed by Braun and Clarke (2006) that consists of data familiarisation, initial coding, theme generation, theme review, theme definition and final reporting. This part explains in detail how the synthesis was operationalised.

The first stage of the analysis consisted of line-by-line reading of the 21 studies included whereby 87 initial codes were inductively generated based on the extracted results. These codes were descriptive, and they remained close to the original language of the studies (e.g., students desire more authentic tasks, instructors complain about lack of class time, administrators do not have any mechanisms of collaboration).

Theme generation was done via iterative deductive and inductive reasoning. The original coding was mainly inductive and the themes that came out of the data were not subject to any predetermined categories. But with patterns starting to emerge, the process became more deductive, with new themes being narrowed down and validated through the entire dataset. This mixed methodology made sure that the resulting thematic framework was data-driven and systematically structured. The curriculum development framework by Richards (2017) was applied to the theme generation only after preliminary theme generation to give theoretical coherence but not a priori categories.

Each of 21 studies covered several themes. No study was assigned a single code. Rather, every study added to several thematic categories based on the scope of the stakeholder perceptions reported. As an

example, one study would include results coded to Curricular Misalignment (students indicating a mismatch between EAP and disciplinary writing), Teacher Expertise Gap (instructors indicating insufficient training), and Institutional Constraints (administrators indicating resource constraints). In those instances, each individual finding was coded individually and the study was matched in the respective themes. This methodology allowed capturing the full depth of the stakeholder perceptions without imposing studies under one thematic label.

Themes (general, repetitive patterns of meaning in two or more studies) and sub-themes (narrower dimensions of a theme that describe particular manifestations or stakeholder orientation) were differentiated. A theme was considered a pattern that emerged in a minimum of three studies and reflected a consistent dimension of stakeholder perception pertaining to EAP curriculum design, delivery, or evaluation. A sub-theme was considered to be a particular aspect, variant or stakeholder-group-specific instantiation of a larger theme. As an illustration, the Curricular Misalignment (Theme 2) had sub-themes such as Genre Mismatch (2.1), Skill-transfer Failure (2.2), Authenticity Deficit (2.3), Audience Awareness Gap (2.4), and Temporal Disconnect (2.5). This hierarchical design enabled the synthesis to work at various levels of granularity.

Theme consolidation was directly impacted by the independent audit done by the second researcher (which is covered in the coding reliability section). The second researcher did not go through each assignment of the code alone, but also through the general structure in terms of the themes during the audit. Where the second researcher found overlapping or ambiguous thematic boundaries, they were noted to be discussed.

Quantitative data was analyzed descriptively if it was possible. Studies with the usage of surveys and rating scales were transferred into a qualitative synthesis. According to Hong (2018), as it was in mixed method reviews, quantitative evidence could be kept with the help of descriptive synthesis without destruction of academic integrity. As a result, this analytical review produced a context-sensitive, comprehensive, and transparent synthesis of the stakeholders' perception regarding EAP curricula in the higher education context.

3. Findings and Discussion

3.1 Overview of Included Studies

The final sample number consisted of 21 empirical studies published between 2000 and August 2025. Each peer-reviewed article examined stakeholders' perceptions regarding EAP curricula in the context of higher education. All the studies were conducted across 12 countries and regions of the world, showing the immense expansion of EAP and different institutional models where it was used. The most concentrated research was conducted in East Asian territories such as China, where the importance of EAP is stated at the national level (Goa & Bartlett, 2014). Other studies were conducted in Hong Kong, Taiwan, Japan, Singapore, Sri Lanka, Thailand, Indonesia, Iran, the Czech Republic, and Canada, as well as international branch campuses operating within China.

When it comes to the methodological orientation, most studies employed qualitative designs, such as semi-structured interviews, focus group discussions, document analysis and thematic analysis, followed by Braun & Clarke's framework. This type of approach usage once again represents the exploratory nature of perception-based research. There were some studies which implemented a mixed-method approach, especially those focusing on needs analysis, which triangulated data from faculty surveys, questionnaires, course outlines and assessment mappings. Even fewer studies adopted purely quantitative survey designs. However, even those surveys capture quantitative surveys with open-ended questions to gather a richer viewpoint. Overall, it was possible to see that most studies demonstrated methodological robustness as they were using data triangulation, member checking or iterative coding to ensure reliability and trustworthiness of the results.

Across the studies, there were the following stakeholder groups: undergraduate and postgraduate students enrolled in English-medium or bilingual programmes; EAP instructors and English language lecturers; subject-specialists teaching content courses; pre-service teachers preparing for future EAP or EFL roles; curriculum designers, EAP coordinators, or institutional administrators.

The largest stakeholder group were students. Their perspectives were around academic challenges, the relevance of EAP to real coursework, and the scope of how EAP could prepare them for discipline-specific demands. The next group in number was instructors and faculty members. They mainly had perspectives regarding teacher preparedness, pedagogical challenges, material design and institutional constraints. The smallest group was related to the administrators and curriculum designers who mainly had perspectives about policy of higher education, curriculum changes and organizational expectations of HEIs. All together those insights create a multi-layered picture of how EAP curricula is experienced in different contexts.

3.2 Thematic Findings

3.2.1 Stakeholder Perceptions of EAP Relevance and Curriculum Design

One concept that was mentioned across studies is the growing significance of EAP in higher education because of internationalization and rising demand for academic literacy. Generally, students see it as highly relevant, especially in writing or research-oriented programs. Pre-service teachers in Hong Kong and China mentioned that traditional College English did not prepare students for academic tasks. Mainly because he was not filling the gap which was between general English and discipline-oriented academic skills. The same view was mentioned by faculty members, adding that EAP is crucial for academic argumentation, source integration and academic integrity.

However, the relevance of EAP was different in different institutional settings. In Western universities like Canada, EAP is often seen as a transitional tool supporting students' entry into the academic university context, whereas the Asian context sees EAP as a curriculum reform agenda demanding shifts in pedagogy, teacher identity and priorities of institutions. Therefore, the design of EAP curricula presented structural and stakeholders' evolving expectations.

3.2.2 Perceived Strengths and Gaps in EAP Delivery

Some studies mentioned strong sides of EAP curricula, such as improved academic writing skills, stronger confidence in summarising, paraphrasing and citation usage. Students usually prefer explicit instructions in language, like grammar accuracy or sentence structure. Especially, from the context where the schooling system was focused on test-oriented learning and memorization.

However, even though there were a lot of strong sides. It was possible to see a lot of weaknesses. The list includes not enough connection between writing assignments and disciplinary tasks, limited teacher preparedness, lack of institutional support, gaps and difficulties in material design, and limited incorporation of academic integrity, digital literacy and research skills. The cases of these weaknesses were due to different reasons and in different contexts. The strongest reason was the transition of teachers from the general English course to the EAP teaching context or a combination of EAP and general English, thus designing English for General Academic Purposes. Other reasons included heavy workload, very limited or no professional development activities, old or generalized textbooks, and inconsistent incorporation of academic integrity, digital literacy and research skills.

In some cases, the absence of any documentation or policy regarding the module became one more fragmentation in curriculum design and delivery.

3.2.3 Curriculum Alignment with Academic and Disciplinary Needs

The biggest misalignment was between EAP curricula and the actual writing, reading and communication demands of the degree programs. Needs analysis in Iran, Sri Lanka, and Canada showed the difference between what students expected from the course and what was actually taught during the

program. While studying disciplinary courses, students were expected to submit case studies, technical reports, abstracts, design proposals and summaries. However, this was not taught during the EAP program.

Also, there was a very strong mismatch between writing-focused fields such as Business, Engineering, Health Studies, and Media & Design and programs which were audience-oriented and expected to have professionally relevant genres. If the first was expecting an academic essay, the second one rarely used it and suggested English for Specific Academic purposes or any hybrid approach which could help to combine both concepts. Additionally, instructors mentioned that there was not enough collaboration between language and content instructors, which caused less discipline-oriented curricula development.

3.2.4 Institutional and Pedagogical Factors

The most decisive role among stakeholders belonged to institutional structuring. Countries working with centralized systems, like China, mentioned the importance of national policy change impact on the development of curriculum development framework, exam-driven EAP teaching, and fixed curriculum frameworks. Also, there were mentioned challenges like low institutional recognition, no training opportunities and limited autonomy in material development for the course. The Canadian and European context in comparison was more supportive, providing structured professional development training, specialized curricula development committees and better developed learning outcomes. However, there were difficulties in balancing EAP with the specialized needs of the stakeholders.

Many studies showed that there was specific tension between the traditional teacher-centred teaching approach and the student-centred interactive approach which is usually recommended in EAP. In most cases, students who studied in the system with passive learning were resistant to modules which required critical thinking implementation, project-based learning, or collaborative writing. In the same way, teachers who had lack of training or institutional help could not implement a new role in the classes. As a result, it affected overall EAP delivery and stakeholders' perceptions of curriculum quality.

3.3 Comparative Insights

3.3.1 Contrasts Between Stakeholder Groups

Different stakeholder groups usually had different evaluations regarding the EAP evaluation. Students, while expressing their perceptions, were focusing on their immediate academic needs. Those could be learning how to write reports, essays, summaries and citations of the work. Instructors were providing insights regarding pedagogical factors in EAP teaching, providing examples regarding the large size of groups, lack of resources, weak feasibility in the disciplinary knowledge, and insufficient preparation for EAP teaching. Content instructors were concerned regarding misalignment among the disciplines which were taught. They stated that EAP courses have limited integration of writing, connection to the specialized field. Finally, administrators and curriculum designers often look at EAP as a module which must focus on institutional goals, which will ensure internationalization, accreditation and EMI expansion at the university.

These various perspectives once again presented the need for coordinated, multi-stakeholder-oriented development of curriculum, which will be aligned, sustainable and relevant for most stakeholder groups.

3.3.2 Contextual Differences (Regional and Institutional)

The differences in perspectives were not only among stakeholders, it varied among different contexts. Western Institutions in Canada and European countries focused their priority on academic integrity, development of research skills, workplace-oriented topics and vocational program demands. However, East Asian universities struggled with changing from general English to academic one and also transitioning from exam-oriented classes with more student-centred hands-on practices. One more difficulty faced was the lack of training for teachers. The Middle East and South Asian contexts provided concerns regarding the importance of aligning EAP with disciplinary tasks the same way mentioning the lack of resources and limited teacher training. One more context was related to joint

venture universities in China, which provided perspectives regarding hybrid practices which combined Western-style academic literacy with local contextual input.

There were different perspectives in different types of universities as well. For example, research-oriented universities were focusing on academic writing and research skills. However, teaching-oriented institutions preferred professional communication, audience awareness. This contextual difference shows that EAP must be implemented considering local educational culture, policy and discipline to make it more stakeholder-oriented and needs-focused.

3.4. Discussion

The findings of this systematic literature review present the complex and developing nature of EAP curriculum implementation in different HEIs. Most stakeholders mentioned the importance of EAP for supporting students' academic success. However, it was possible to see huge gaps in curricula alignment, teacher readiness, and institutional support. The following discussion will interpret the findings, including global and local trends.

3.4.1 Interpretation of the Main Findings

In internalization-focused and English medium institutions, EAP is generally perceived as a special program which prepares students for academic success. Students in different contexts stated the relevance of EAP for writing, reading and communication skills enhancement. It was mostly used to help learners navigate disciplinary courses. This idea is aligned with scholars indicating that academic literacy is central to students' success in higher educational contexts (McWilliams & Allan, 2014; Murray, 2022). However, there was a mismatch between the EAP curricula that were taught and the actual academic tasks which students encountered during their academic journey. Especially, this was clearly seen in the EGAP modules and less vivid in ESAP courses.

This tension was also supported by teachers' perceptions. Many EAP instructors reported that they felt underprepared for the course teaching as their training background was oriented toward general English teaching and lacked academic literacy or discipline-oriented pedagogy. That was one more example provided by Liu and Zou (2021), regarding the shift from language-focused instructions to academic ones. Additionally, instructors from different regions mentioned institutional limitations such as heavy teaching loads, limited access to professional development and limited curricula independence. All those factors were a big obstacle in EAP curriculum development.

It is important to mention policy-level perspectives as well. As an example, in China, the nationwide reform regarding EAP lifted its status. However, there was strong pressure on curriculum development and transformation. As it had no proper support for teacher training or resource development. In the Western context, the development of EAP is more flexible, but there were challenges regarding the balance between academic literacy and discipline-oriented expectations. All the institutions perceived EAP as a bridge that connects institutional, pedagogical, and disciplinary priorities.

3.4.2 Implications for Curriculum Development and Institutional Policy

The result provides several solutions that must be applied to strengthen the EAP provision. Initial is misalignment between EAP curricula and disciplinary requirements. This challenge highlights the importance of Needs analysis as the most important pillar of curriculum design. In order to overcome the challenge, institutions must map assessment practices, communication expectations, and disciplinary demands between various fields. Collaboratively designed curriculum involving all stakeholders can connect disciplinary and language lacks and enhance curricula coherence.

The second important action requires professional development, which is tailored to the EAP teachers. EAP demands knowledge in academic integrity, digital literacy, genre-oriented and research skills; therefore, training that will provide a smooth transition from general English to EAP would facilitate and improve teachers' competency. Training could include workshops, co-teaching models, and institutional support for research. All those topics would make the shift process a more structured way

The third point is that EAP should be recognized as a strategic subject which helps with internationalization, student mobility and program quality assurance. This recognition of EAP would be followed by greater investment into professional development, materials development and better working conditions. These reforms will integrate EAP as a subject which provides literacy for further studies rather than a stand-alone subject which is core but not needed.

Finally, Curriculum development will be improved if authentic, task-based, and genre-oriented pedagogy is implemented. Especially in the EMI setting. As all those factors prepare for better results in further studies and the preparedness of students to disciplinary genres.

3.4.3 Alignment with Global Trends and Local Educational Realities

One more factor that was identified in the articles was the tension between global academic norms and local educational constraints. On one hand, global trends like the international ranking system, the expansion of EMI and partnerships among universities increased the importance of English language proficiency and communication skills. On the other hand, local realities make the whole process challenging to unfold. In most Asian and Middle Eastern contexts exam exam-oriented culture is in priority. Most universities are following hierarchical institutional structures; as a result, teacher autonomy is very limited, which prevents innovation and progress. Local teacher training systems, resources and sociolinguistic norms shape expectations of stakeholders of how global practices in EAP could be adapted. Considering all these factors, EAP development must be smooth without rapid transitions. It must be context sensitive that acknowledging international standards of the educational system of the particular region. Thus, learners are better prepared to the academic life and quality education is supported with the help of integration and implementation of those skills from the first year of education in the institutions.

4. Recommendations

4.1 Suggestions for EAP Curriculum Enhancement

The results showed that EAP curricula will be more responsive to the actual academic and disciplinary needs of students (backed by 18 out of 21 studies; very high evidence strength). This might be done through the inclusion of routine needs analysis that is critical to afford alignment between evaluation, disciplinary needs, and communicative practices to which students are subjected in their programs (18/21 studies; very high). Connecting authentic academic tasks would by integrating English for Specific Academic Purposes (ESAP) clearly demonstrate the relationship between what students learn and what students need during the course supported with 17 studies (high evidence strength).

Another significant aspect, which should be considered, is that EAP has to be instructed more than general academic skills (15/21 studies; high). Literacy skills, critical thinking, multimodal literacies, academic integrity and problem-solving skills are vital components that should be incorporated in globalized higher education (academic integrity: 16/21 studies, high; multimodal literacies: 12/21 studies, moderate).

Moreover, the review also points to the critical importance of stakeholder partnership in terms of the improvement of EAP curricula (18/21 studies; very high). Collaboration between all the involved stakeholders (EAP instructors, content instructors, program coordinators, policy and decision makers (cross-departmental committees: 14/21 studies, high; co-teaching models: 11/21 studies, moderate) needs to be established. Co-teaching in practice, the formation of cross-departmental curriculum planning committees and development of common assessment might form a coherent academic literacy improvement environment.

The primary recommendation is the one about the investment in professional development of EAP instructors who are oriented on genre pedagogy, academic writing, corpus-informed teaching, and disciplinary knowledge development (genre pedagogy training: 19/21 studies, very high; corpus-informed teaching: 10/21 studies, moderate; disciplinary knowledge development: 17/21 studies, high).

The detailed recommendations retrieved from the studies and collected from the analysis of the papers is provided in Table 2.

Table 2. Evidence grading box for recommendations

Recommendation	Specific Action Proposed	Strength of Evidence (Studies Supporting)
Needs-Driven Curriculum Design	Conduct regular, systematic needs analyses involving students and content faculty	Very High (18/21)
	Align EAP tasks with actual disciplinary assignments (reports, case studies, proposals)	Very High (19/21)
Disciplinary Integration & Authenticity	Integrate English for Specific Academic Purposes (ESAP) alongside general EAP	High (17/21)
	Use authentic materials from students' disciplines (research articles, technical reports)	High (16/21)
Expansion of Academic Literacies	Teach critical thinking and source evaluation alongside language skills	High (15/21)
	Integrate digital and multimodal literacy skills	Moderate (12/2)
	Explicitly teach academic integrity and ethical scholarship practices	High (16/21)
Stakeholder Collaboration & Integration	Establish formal collaboration mechanisms between EAP and content faculty	Very High (18/21)
	Create cross-departmental curriculum committees	High (14/21)
	Implement co-teaching or team-teaching models	Moderate (11/21)
Teacher Professional Development	Provide genre pedagogy and disciplinary writing training for EAP instructors	Very High (19/21)
	Offer corpus-informed teaching and materials development training	Moderate (10/21)
	Support EAP teachers' disciplinary knowledge development	High (17/21)
6. Institutional Support & Recognition	Improve EAP instructors' status and workload conditions	High (16/21)
	Integrate EAP as credit-bearing courses within degree programs	Moderate (13/21)
	Provide adequate resources (class size, materials, technology)	Very High (18/21)
7. Sustainability Integration	Incorporate sustainability-focused reading and research materials	Emergent (5/21)
	Design project-based assignments addressing sustainability topics	Emergent (4/21)
	Foster interdisciplinary collaboration around sustainability themes	Emergent (3/21)
8. Pedagogical Innovation	Shift from teacher-centered to student-centered, interactive approaches	High (17/21)
	Implement task-based and project-based learning methods	Moderate (14/21)
	Provide more speaking and interactive practice opportunities	High (16/21)

Despite not being an analytical claim of this review, a few of the included studies (n=5) expressed the possibility of including sustainability competencies in EAP curricula. These are some of the new results that may indicate that the reading materials, data analysis activities, and writing of research about the environment or social issues might be used as both the means of building academic literacy and awareness of world issues. Project-based assignments based on sustainability-related topics, and

interdisciplinary cooperation between language teachers and specialists in the subject and working on sustainability-related content were also suggested in some studies.

The findings in the Systematic Literature Review show important areas for improvement in EAP such as curriculum design, delivery and policy. Provided recommendations are given to support EAP curriculum designers, administrators and institutional governance and higher educational policy makers to make effective decisions. Especially, they are focused on needs-analysis framework development, designing structured and targeted professional development, establishment of specific mechanisms for content and language instructors' mechanisms.

4.2 Directions for Future Research

The future studies should extend the geographical area of EAP curriculum research to Central Asia, Africa, Latin America and Eastern Europe since cross-institutional research with a comparative mode would provide more information about the structural, cultural and linguistic background that influences the perception of stakeholders. One more suggestion might be long-term research since it would show not only the process of reform and transition but also what impact has been obtained and what results could be achieved. Besides, studies that combine multiple types of data, including document analysis, classroom observations, and corpus-informed studies would offer a more holistic and comprehensive view of EAP curriculum efficacy. In particular, the emergent theme of sustainability integration found in this review needs to be systematically investigated. Future research ought to explore the feasibility and ways in which sustainability issues can be sincerely incorporated into EAP without compromising disciplinary language objectives, effectiveness of sustainability-centric EAP activities on both academic literacy products and sustainability consciousness, and cross-cultural variations in EAP programmes sustainability approaches, especially in under-researched areas. When we reflect on all the suggestions, one should keep in mind that in case of conclusions, one should be cautious about the possibility to generalise since the literature reviewed was limited both geographically and methodologically. The researches have predominantly focused on East Asia and the western countries leaving a huge gap to other regions like central Asia, Africa and Latin America. Moreover, quality qualitative methodologies were able to give in-depth details of particular situations and restrict generalisation. As such, the recommendations in Table 2, particularly the emergent sustainability theme cannot be viewed as universal prescriptions.

It should be noted that prior to making cross-contextual comparisons it is critical to note that there are numerous possible biases that can affect the perceived regional trends that are reported in this review. First, the inclusion criterion of studies limited to English-language peer-reviewed journals can have resulted in a publication bias because any studies with negative findings, or any studies published in non-English journals, or any studies based on non-Western education systems and thus having different publication cultures might be underrepresented. Second, institutional reporting cultures can differ significantly between regions; e.g. the higher education systems in East Asia might have a higher number of norms and incentives to publish curriculum evaluation research than Western ones, which may influence the amount and nature of the evidence available. Third, the high rate of East Asian and Western studies in the literature reviewed suggests that the trends in these areas could be stronger, whereas the results of Central Asia, Africa, Latin America, and Eastern Europe could be viewed as provisional because of the lack of representation. This is to encourage readers to take cross-contextual comparisons with due caution, as differences or similarities seeming to be present may be as much due to reporting bias as to actual educational phenomenon

References

- Adarkwah, M. A., & Zeyuan, Y. (2020). Perceptions of the Effect of an EAP Course on English Self-Efficacy and English Proficiency: Voices of International Students in China. *Online Submission*, 10(2), 1-17.
- Afshar, H. S., & Movassagh, H. (2016). EAP education in Iran: Where does the problem lie? Where are we heading?. *Journal of English for Academic Purposes*, 22, 132-151
- Aksit, T., & Aksit, N. (2022). Establishing an institutional EAP teacher development scheme based on BALEAP's competency framework: A critical review of the competencies. *Journal of English for Academic Purposes*, 60, 101175.
- Al Issaei, N. S. A. (2018). *Investigating the perceptions of EAP teachers about Curriculum challenges in colleges of Applied Sciences in Oman*. University of Exeter (United Kingdom)
- Al-Zubidy, A., & Carver, J. C. (2019). Identification and prioritization of SLR search tool requirements: an SLR and a survey. *Empirical Software Engineering*, 24(1), 139-169.
- Azarian, M., Yu, H., Shiferaw, A. T., & Stevik, T. K. (2023). Do we perform systematic literature review right? A scientific mapping and methodological assessment. *Logistics*, 7(4), 89.
- Azarian, M., Yu, H., Shiferaw, A. T., & Stevik, T. K. (2023). Do we perform systematic literature review right? A scientific mapping and methodological assessment. *Logistics*, 7(4), 89.
- Bakogiannis, A. (2025). Exploring Inclusive Teaching Practices of English for Academic Purposes (EAP) in Higher Education (HE): Progress, Challenges and Next Steps. *The Language Scholar*, 7.
- Bo, W. V., Lim, L., & Lin, Y. (2022). An investigation of multilingual domestic university student perceptions of English for academic purposes. *Behavioral Sciences*, 12(9), 328.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Carroll, C., & Booth, A. (2015). Quality assessment of qualitative evidence for systematic review and synthesis: is it meaningful, and if so, how should it be performed?. *Research synthesis methods*, 6(2), 149-154.
- Csiernik, R., Cavell, M., & Csiernik, B. (2021). EAP evaluation 2010–2019: What do we now know?. *Journal of Workplace Behavioral Health*, 36(2), 105-124.
- Ding, A., & Bruce, I. (2017). *The English for academic purposes practitioner*. London: Springer, 10, 978-3.
- Dissanayake, K. M., & Harun, R. N. S. R. (2012). Theory and practice of EAP in the Sri Lankan context. *Procedia-Social and Behavioral Sciences*, 66, 106-116
- Dooley, P. (2010). Students' perspectives of an EAP pathway program. *Journal of English for academic purposes*, 9(3), 184-197
- Dou, A. (2024). Needs analysis of ESP courses from the perspective of students and teachers. *Frontiers in Educational Research*, 7(6), 132-138.

- Douglas, S. R., & Kim, M. (2014). Task-based language teaching and English for academic purposes: An investigation into instructor perceptions and practice in the Canadian context. *TESL Canada Journal*, 1-1.
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6(1), 3-17.
- Fendt, J. (2025). Embracing emergence in qualitative meta-analysis: A guide to higher-order synthesis. *Methodological Innovations*, 18(3), 143-167.
- Floyd, C. B. (2015). Closing the gap: International student pathways, academic performance and academic acculturation. *Journal of Academic Language and Learning*, 9(2), A1-A18.
- Galloway, N., Kriukow, J., & Numajiri, T. (2017). Internationalisation, higher education and the growing demand for English: An investigation into the English medium of instruction (EMI) movement in China and Japan.
- Gao, Y., & Bartlett, B. (2014). Opportunities and challenges for negotiating appropriate EAP practices in China. *English for Academic Purposes (EAP) in Asia: Negotiating appropriate practices in a global context*, 13-31.
- Harper, J., & Sun, Y. (2022). EAP Courses in joint-venture institutions: A needs analysis based on learner perceptions. *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)*, 7(1), 159-179
- Helmer, K. A. (2013). Critical English for academic purposes: Building on learner, teacher, and program strengths. *Journal of English for Academic Purposes*, 12(4), 273-287
- Hessamy, G., & Mohebi, M. (2014). Perceptions of Medical Students and EFL Instructors of Their EAP Materials, Challenges and Implications for Iranian EAP Instructors. *Journal of Language Teaching & Research*, 5(1)
- Hong, Q. N. (2018). *Revision of the Mixed Methods Appraisal Tool (MMAT): A mixed methods study*. McGill University (Canada).
- Hu, G., & Lei, J. (2014). English-medium instruction in Chinese higher education: A case study. *Higher education*, 67(5), 551-567.
- Huang, J. C. (2024). Unveiling EFL graduate students' EAP needs and perceptions of EAP courses: A large-scale survey in Taiwan. *Journal of English for Academic Purposes*
- Hyland, K. (2006). The 'other' English: thoughts on EAP and academic writing. *The European English Messenger*, 15(2), 34-38.
- Hyland, K. (2016). General and specific EAP. In *The Routledge handbook of English for academic purposes* (pp. 17-29). Routledge.
- Iranmehr, A., Atai, M. R., & Babaii, E. (2018). Evaluation of EAP programs in Iran: Document analysis and expert perspectives. *Applied Research on English Language*, 7(2), 171-194.
- Jančaříková, R., Povolná, R., Dontcheva-Navratilová, O., Hanušová, S., & Němec, M. (2020). An academic writing needs analysis of Czech university graduate students. *Discourse and Interaction*, 13(1), 42-66.

- Joanna Briggs Institute. (2017). The Joanna Briggs Institute critical appraisal tools for use in JBI systematic reviews: checklist for prevalence studies. *Retrieved November, 15, 2018.*
- Kanokpermpoon, M. (2023). Foundation English Course Development for Thai Tertiary Education: Learners' Perceived Needs and Factor Analysis. *REFlections*, 30(3), 671-694.
- Keefe, K. (2017). An EAP program and students' success at a Canadian university. *TESL Canada Journal*
- Khany, R., & Tarlani-Aliabadi, H. (2016). Studying power relations in an academic setting: Teachers' and students' perceptions of EAP classes in Iran. *Journal of English for Academic Purposes*
- Kirkpatrick, D. L. (1998). The four levels of evaluation. In *Evaluating corporate training: Models and issues* (pp. 95-112). Dordrecht: Springer Netherlands.
- Kohnke, L., Zou, D., & Zhang, R. (2022). Exploring pre-service teachers' beliefs about EAP instruction in China—A case study. *International Journal of Educational Research Open*, 3, 1001
- Lu, X., & Zou, B. (2021). Academic teachers' perceptions of content and language integrated learning (CLIL) programme with EAP teachers. *Pedagogies in English for academic purposes: Teaching and learning in international contexts*, 123-135.
- McWilliams, R., & Allan, Q. (2014). Embedding academic literacy skills: Towards a best practice model. *Journal of university teaching and learning practice*, 11(3), 1-20.
- Murray, N. (2022). A model to support the equitable development of academic literacy in institutions of higher education. *Journal of further and Higher Education*, 46(8), 1054-1065.
- Öğretim, S. B. A. A. İ. (2020). Towards a Defensible EAP Curriculum: A Needs Analysis. *Journal of Theoretical Educational Science*, 13(2), 424-439.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *Journal of clinical epidemiology*, 134, 103-112.
- Pomat, N. (2025). Updating the Landscape: A Systematic Review of Emerging Trends in English for Academic Purposes (EAP) Research. *Language Teaching Research Quarterly*, 47, 255-280.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version, 1*(1), b92.
- Richards, J. C. (2001). *Curriculum development in language teaching*. Cambridge University Press.
- Sloan, D., & Porter, E. (2010). Changing international student and business staff perceptions of in-session EAP: using the CEM model. *Journal of English for Academic Purposes*
- Smith, V., Windle, S., & Johnny, L. (2025). Meeting the needs of EAP students in Canadian colleges: Insights from a needs analysis study. *English for Specific Purposes*, 78, 139-155.
- Song, Y., & Zhou, J. (2022). Revising English language course curriculum among graduate students: An EAP needs analysis study. *Sage Open*, 12(3).
- Tajino, A., James, R., & Kijima, K. (2005). Beyond needs analysis: soft systems methodology for meaningful collaboration in EAP course design. *Journal of English for Academic Purposes*

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC medical research methodology*, 8(1), 45.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International journal of bilingual education and bilingualism*, 9(2), 159-180.

Appendices

Appendix I: Searching Strings for database

1. Scopus:

Search date: 13/05/2025

Exact Search String:

TITLE-ABS-KEY (("Perceptions" OR "Perspectives" OR "Attitudes") AND ("Stakeholders" OR "Teachers" OR "Students" OR "Administrators") AND ("EAP" OR "English for Academic Purposes") AND ("Curriculum" OR "Syllabus" OR "Program" OR "Course") AND ("Higher Education" OR "University")) AND PUBYEAR > 1999 AND PUBYEAR < 2025

Filters Applied: Article Title , Abstract, Keywords; Publication Years 2000-2025; English Language; Article

2. ERIC

Search Date: 27/05/2025

Exact Search String:

(“Perceptions” OR “Perspectives” OR “Attitudes”) AND (“Stakeholders” OR “Teachers” OR “Students” OR “Administrators”) AND (“EAP” OR “English for Academic Purposes”) AND (“Curriculum” OR “Syllabus” OR “Program” OR “Course”) AND (“Higher Education” OR “University”)

Filters applied: Peer reviewed; Publication Date: Since 2000; Journal Articles

3. ScienceDirect:

Search Date: 09/06/2025

Exact Search String:

("Perceptions" OR "Perspectives") AND ("Teachers" OR "Students" OR "Administrators") AND ("English for Academic Purposes") AND ("Curriculum" OR "Syllabus" OR "Program")

Filters applied: Years 2000-2025; Article type; Research Article

4. SpringerLink:

Search Date: 17/06/2025

Exact Search String:

("Perceptions" OR "Perspectives" OR "Attitudes") AND ("Stakeholders" OR "Teachers" OR "Students" OR "Administrators") AND ("EAP" OR "English for Academic Purposes") AND ("Curriculum" OR "Syllabus" OR "Program" OR "Course") ("Higher Education" OR "University")

Filter Applied: Content type: Research Article; Date Published:2000-2025; Languages: English; Subjects: Language Teaching and Learning, Higher education,

Appendix II:

Citation	Country	Sampled stakeholder group, size	Focus	score
Douglas & Kim (2014)	Canada	46 students (Chinese, Indonesian)	learner perceptions of course content, teaching methods, and relevance to academic tasks	6
Keefe (2017)	Canada	8 international students (7 China, 1 Korea)	Perception of 8 week intensive EAP program	10
Hurper & Sun (2022)	China	16 undergraduate Chinese students	Perseption of students on one year programm	10
Kanokpermpoon (2023)	Thailand	145 first-year Thai EFL students	Foundation English (FE) course; assessed students' perceived needs in reading, writing, listening, speaking, and academic study skills	10
Öğretim, S. B. A. A. İ. (2020).	Turkey	41 instructors in 5 focus groups; 37 student representatives in workshops; 61 repeat students completing open-ended questionnaires	Needs analysis (NA) study and subsequent curriculum innovation; phases: 1) Informal meetings with administrative bodies, 2) NA research through instructor focus groups and student workshops, 3) Evaluation via open-ended questionnaires; innovations included skill-based activities, portfolio writing, project grades (PG) instead of class performance grades (CPG), optional	10
Song, & Zhou(2022)	China	200 graduate students, 4 instructors	EAP skills in academic listening, speaking, reading, writing; questionnaire + semi-structured interviews	12
Afshar, H. S., & Movassagh, H. (2016)	Iran	831 B.A. students; 55 university teachers; 40 students and 25 teachers (subset) interviewed; 35 EAP classes observed; 30 EAP syllabi analyzed	EAP course relevance, classroom observations. EAP focus	12
Dissanayake & Harun (2012)	Sri Lanka	30 first-year undergraduates,	EAP needs in relation to academic English: language structures, rhetorical categories, language functions, language skills (listening, speaking, reading, writing), study skills	9
Dooley, P. (2010)	Australia	13 international students enrolled in the English Language	4 ELBC units: Academic Writing, English Communication, English for Specific Purposes	8

		Bridging Course (ELBC).	(ESP), Tertiary Study Skills. Two rounds of semi-structured interviews	
Evans, S., & Green, C. (2007)	China	4932 undergraduates	Large-scale needs analysis using: • Self-report questionnaire on writing, reading, speaking, listening • Interviews with students & lecturers • Focus groups with programme leaders • Writing & speaking tests (4 stages)	12
Helmer, (2013)	The United States of America	Students 121; Faculty 6; Writing-center tutors 7	Critical EAP (CEAP) needs analysis of a 10-year-old EAP program inside a larger Writing Program. Analysis included surveys, focus groups, classroom observations, portfolio review. Goal: understand institutional structures, align curriculum, and redesign the program.	10
Huang, J. C. (2024)	Taiwan	1155 graduate students; 40 student interviewees; 22 professors	Needs analysis of EAP reading/writing, Students' perceived needs in reading/writing; Professors' perceptions of students' EAP skills; course expectations and satisfactions	10
Khany & Tarlani-Aliabadi (2016)	Iran	Students: 200; Teachers: 50 EAP Interview subset: 18 students + 10 teachers.	Exploration of the extent to which students' and teachers' rights, feedback, and voices are incorporated into EAP curriculum decisions. • Focus on textbook selection, teaching methods, assessment practices, participation patterns, course objectives, and classroom environment. • Critical EAP framework (Benesch, Freire, Foucault).	9
Sloan & Porter (2010)	The United Kingdom	150 Students; 7 program directors	Collaboration to align EAP with disciplinary aims; Development of materials based on discipline-specific genres; Curriculum alignment (constructive alignment principles); Evaluation of existing EAP provision in business school context	11
Tajino, James, & Kijima (2005)	Japan	29 students	Soft system methodology application in needs assessment and perception analysis	12
Adarkwah, & Zeyuan. (2020)	China	8 students	Course designed focusing on the improvement of academic skills of low level students	12
Bo, Lim, & Lin, (2022)	Singapore	50 students	Perceptions regarding all academic skills development during the course	12
Hessamy, & Mohebi (2014)	Iran	175 students, 75 EFL instructors	Focus on book aspects, subject matter, tool usage	6
Jančaříková & et. al (2020)	Czech	67 ELTE students	Comprehension of academic writing needs analysis	6

Kohnke, Zou, & Zhang(2022)	China	16 Students	Study explored preservice teachers' perceptions of the existing provision of EAP in mainland China and the challenges and opportunities for implementing EAP	6
Smith, Windle, & Johnny (2025)	Canada	44 enrolled post secondary course, focus group 11students and 7 professors	Needs Analysis (NA) to understand what writing assignments and skills are required in Programs of Study (POS) and how well the EAP program prepares students.	6

Assessing Knowledge, Attitude, And Practice (KAP) in Zakat Management Among Higher Education B40 Students

ABSTRACT

Authors: Adibah Bahori¹, Izuli Dzulkffli², Fathiyah Mohd Fakhruddin², Suriani Ismail²

Affiliation: ¹Academy of Contemporary Islamic Studies, Universiti Teknologi MARA, 71300 Rembau, Negeri Sembilan, Malaysia
²Faculty of Educational Studies, Universiti Putra Malaysia

*Corresponding author:
izuli@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026:
| Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

B40 students come from financially challenged backgrounds or from low-income families. This group is often targeted for government aid and support programs such as zakat, especially in education. This study aims to assess Knowledge, Attitude, and Practice (KAP) among B40 students regarding zakat management at UiTM Negeri Sembilan. A total of 236 B40 students participated through a quantitative approach using a Likert-scale questionnaire. Descriptive and correlation analyses were conducted using SPSS. The findings indicate that KAP levels are high, with most mean scores exceeding 4.5, reflecting strong awareness of zakat application procedures, positive attitudes toward its effectiveness in reducing financial burdens, and consistent practices in complying with application requirements. Students also expressed satisfaction with the user-friendly e-zakat system and clear communication channels. These findings can guide zakat institutions to enhance university policies and develop communication strategies that align with students' knowledge levels, attitudes, and practices. This study provides valuable insights that can benefit various stakeholders, including zakat management institutions, UiTM students, and policymakers to help B40 students to get quality education in their studies. By addressing communication gaps and enhancing student awareness and engagement, the study aims to create a more inclusive and efficient zakat distribution system, ultimately aiding students from the B40 category in their financial and educational pursuits.

Keywords: B40 Students, Higher Education, Knowledge, Attitude, and Practice (KAP), Zakat, Zakat Management.

1. Introduction

Zakat helps ease the financial burden of needy individuals, including B40 students. B40 students refer to individuals who come from families within the bottom 40% income group in Malaysia. This classification is based on monthly household income and is used to identify those who may require financial assistance (Nor Samsiah et al., 2020). The term B40 stands for “Bottom 40 Percent.” It includes households with a monthly income of less than RM4,850. This group is often the main target for various government aid and support programs, especially in the field of education. In the Malaysian context, zakat recipients within higher education institutions predominantly comprise students from B40 households, where limited household income directly contributes to financial vulnerability and eligibility for zakat assistance.

Students in the B40 category typically come from financially challenged backgrounds. They often face limited access to learning resources such as digital devices, internet connectivity, and academic materials. As a result, they are eligible for various forms of assistance, including scholarships, special financial aid, living allowances, tuition fee waivers, and digital device support. It helps prevent students from low-income families from being left behind in higher education (Mustakim et al., 2024). Indirectly, this effort contributes to social and economic development through education.

The Zakat, Sedekah, and Waqf Unit (ZAWAF) have managed the Zakat Fund for UiTM students in the B40 category since October 2020, utilizing a web-based zakat application to disseminate information about zakat and its distribution process. Despite the provision of zakat assistance for B40 students at UiTM Negeri Sembilan, limited empirical evidence exists on how effectively these students understand, perceive, and engage with the zakat application and distribution system. While previous research has examined aspects of zakat governance and online application mechanisms, there remains a gap in assessing students’ Knowledge, Attitude, and Practice (KAP) within the higher education context. Additionally, it is unclear whether satisfaction toward zakat management varies across demographic factors such as campus location, household income, and academic performance. Therefore, a systematic evaluation of KAP and satisfaction is needed to strengthen zakat governance, communication strategies, and support services for B40 students.

Despite extensive research on zakat management, existing studies predominantly emphasise contributors’ behaviour, institutional governance, or beneficiaries’ understanding. Limited attention has been given to zakat beneficiaries, particularly within higher education institutions, where zakat functions both as financial assistance and developmental support. Moreover, recipient-focused studies often examine awareness or satisfaction in isolation, without integrating behavioural dimensions. Addressing this gap, the present study examines zakat engagement among university students from the beneficiaries’ perspective using the Knowledge–Attitude–Practice (KAP) framework, while modelling Satisfaction as an evaluative outcome of institutional service delivery.

Accordingly, this study focuses specifically on B40 students who directly receive zakat assistance, whose knowledge, attitudes, and practices are shaped by lived beneficiary experiences. This study aims to assess Knowledge, Attitude, and Practice (KAP) among B40 students regarding zakat management at UiTM Negeri Sembilan to get a quality education in their studies.

The research objectives of this study are as follows:

1. To describe the levels of Knowledge, Attitude, Practice (KAP) and Satisfaction toward zakat management among B40 students at UiTM Negeri Sembilan.
2. To determine the difference in practice toward zakat management by campus, household income category and CGPA among B40 students at UiTM Negeri Sembilan.
3. To determine the relationship between Knowledge, Attitude, Practice (KAP) and Satisfaction toward zakat management among B40 students at UiTM Negeri Sembilan.

2. Literature Review

The application of the Knowledge–Attitude–Practice (KAP) framework in this study is grounded in established behavioural scholarship that consistently positions knowledge as a cognitive foundation for attitudinal evaluation, which in turn informs behavioural engagement. Recent empirical studies further demonstrate the relevance of this framework in institutional and service-delivery contexts, supporting its suitability for examining zakat engagement among higher education recipients.

2.1. Financial Education Aid

Financial education aid is a common form of assistance provided by zakat institutions to support needy and underprivileged students from secondary to tertiary levels. Zakat institution is accountable for managing zakat collection and distribution to eight asnaf (Aminuddin et al., 2020). The efficiency and effectiveness of zakat funds collection and distribution are crucial as it is the substance in Islamic economics that can improve the lives of the poor and the needy (Takril & Othman, 2020) known as *Amil*. *Amil* in zakat institutions play a vital role in managing the collection and distribution of zakat funds. Efficient and well-structured distribution systems significantly contribute to poverty alleviation within the community.

Majlis Agama Islam Negeri Sembilan (MAINS) has mobilized special assistance known as the Covid-19 Initiative starting in March to June 2020 amounting to RM8.5 million. In line with that, the source of this special allocation is obtained from the collection of zakat paid by all zakat payers in Negeri Sembilan (Anwar, 2021). Thus, the distribution of zakat payers' property to the asnaf group can ease their burden to buy necessities and increase the source of household income of certain groups as well as move the country's economic growth.

There were studies and writings related to zakat either from the point of view of administration, management, effectiveness of zakat collection and distribution, legal aspects related to zakat, its impact on the economic development of Muslims and so on. The exploration of research related to zakat is very desirable to be developed especially in the face of the current technological challenges that are rapidly developing using digital applications and online zakat are now widely used in facilitating matters including spreading, developing and promoting zakat to the public (Ibrahim, 2015).

Most existing research focuses on the community's perception and acceptance of online zakat payment systems such as a study by Muda et al. (2013) titled Perception of LZS Website Users Towards Online Zakat Payment, Paizin & Som (2019) titled Community Receiver to Pay Zakat Through Online Channels (Online): A New Review and Omar et al. (2022) The Acceptance of Online Zakat Payment Among Muslims in East Malaysia: A Conceptual Study.

Sharif et al. (2022) did a study on the Perceptions on E-Zakat Advancement Distribution Concept Among UiTM Melaka Staff: A Qualitative Study, focusing on UiTM Melaka staff regarding the online zakat distribution mechanism. The study found that most respondents agree on the concept of distribution offered because there was flexibility and freedom in the aspect of distribution given to *muzakkiy* (an individual who is obligated to pay zakat), but some other aspects such as equality in acceptance to the entire asnaf category should be prioritized to ensure that the main goal of zakat is achieved.

The findings of the study also explained how online transfer and distribution methods occur. *Muzakkiy* will transfer the zakat payment directly to the asnaf account whose zakat recipient name was previously screened by *Baitulmal*. In Malaysia, *Baitulmal* refers to an Islamic financial institution responsible for managing and distributing wealth and resources for the welfare of the Muslim community. This transfer follows the same online money transfer concept used by platforms such as Maybank2u, CIMB Clicks, and other similar services. This money transferred concept gives *muzakkiy* the freedom to choose which

type of asnaf to give and gives them the opportunity to make a transfer payment directly to the asnaf's bank account found in the system according to the district and area where they reside and settled.

2.2. Zakat in Higher Education

Studies on zakat at the higher education level have been conducted by several researchers before, especially in the aspect of management and in the aspect of the distribution of zakat to students of higher learning institutions. This shows how the zakat instrument supported the asnaf, helping them improve their standard of living and potentially become future zakat payers (Anwar and Roziah, 2021).

University students are included in the asnaf group, and some studies have discussed it. For example, a study by Hamzah et al. (2021) found that the asnaf students use their zakat money on food and drink expenses, daily needs, and helping the family. Shaharin et al. (2021) stated that the number of students eligible for zakat is increasing due to the COVID-19 pandemic. As a result, zakat assistance to underprivileged students can help ease their burden when studying at university and increase focus on their studies (Aziz et al., 2020).

Hainnuraqma et al. (2024) in their study 'Perceived usefulness and attitude toward intention and acceptance of e-payment zakat' investigate Malaysian Muslims' acceptance of the e-payment zakat system. Five constructs were examined in the survey, based on Technology Acceptance Model, which are Perceived Usefulness, Perceived Ease of Use, Attitude, Intention, and User Acceptability. Meanwhile, this study aims to assess Knowledge, Attitude, and Practice (KAP) among B40 students regarding zakat management at UiTM Negeri Sembilan.

Management of zakat funds in public universities focuses on identifying and supporting needy students (asnaf) via dedicated campus units, often utilizing e-Zakat systems for efficiency. These funds, distributed as monthly or one-off payments, aim to assist with educational expenses and improve academic success (Abdillah et al., 2024).

Abdullah (2017) stated that each Higher Education Institution (HEI) is the administrative place for zakat distribution which can have a significant and positive impact as a one-stop center for students, especially the poor asnaf at Universiti Teknologi MARA (UiTM), Shah Alam. The survey found that most respondents agreed and accepted that this concept was implemented in an educational institution. In addition, the study also proposed a model of administration of zakat distribution in institutions of higher learning (HEI) based on the concept of localization as well as the concept of decentralized power in the aspect of the collection and distribution of zakat.

This view is in line with Muhammad Tajuddin & Wan Mohd Al Faizee's study (2017) stating that the effectiveness of zakat management at Universiti Kebangsaan Malaysia is shared with Lembaga Zakat Selangor (LZS) so that it exists to the best standards in its management. The appointment of an assistant amil from among higher learning institutions is the right approach. The sharing of ideas, work and energy among IPTA/S can further streamline the governance of zakat in the state of Selangor. The UKM zakat unit under the UKM Zakat Trust Fund (TAZUKM) committee has taken on the responsibilities given by LZS well. To date, the zakat collection figures at UKM have shown an impressive increase from year to year.

Some studies found that the online zakat application process has many positive effects in the aspect of zakat distribution and governance by those who manage zakat applications. Applicants among students no longer use forms manually and every application information is easily accessible online through the website provided (Alam, 2023). This efficient distribution of zakat attracted the interest of the Sabah Zakat Center and Etiqa Takaful Berhad to channel zakat funds to UiTM Sabah Branch in helping the asnaf category among students (Izani & Aripin, 2017). The findings of the study also found that online zakat distribution has an effective distribution effect in addition to successfully attracting zakat funds from external parties amounting to RM 560,000.00 for the period Jan - June 2016. Nasrul Hisyam et al

(2018) found that zakat management in universities can ease the burden on students in furthering their studies.

Mat Daud & Wahid (2024) found that financial aid, through zakat indeed increases positive effects in the learning process and zakat aid is found to increase student readiness in the education process and academic performance of undergraduate students. Their study also found that student readiness in the education process plays a significant role in mediating the effect between zakat aid and the learning process and academic performance. This study also demonstrates that the learning process is equally important for student readiness and academic performance.

A study by Azman and Ahmad Wifaq (2011) on the condition of a student as a recipient of zakat is considered through the ruling of distributing zakat to students based on the opinions of contemporary and previous scholars. The results of this study found that there is dissatisfaction from some students who apply for zakat education because they have difficulty in getting help from zakat institutions. In addition, it is identified among the factors that cause students to not be able to get zakat assistance because they do not meet the eligibility requirements that need to be met before the application is made.

In this regard, a study conducted by Nur Azura et al. (2004), found that a high percentage of respondents suggested that UUM administrators hold talks, briefing workshops and courses to create aggressive awareness on income zakat that can be participated. This study focused on the importance of using intermediate mediums in conveying information on income zakat. Studies related to Zakat at HEI were done by Shamsuddin et al. (2021) which studies the impact of money on HEI students during the Covid 19 pandemic. The research found that zakat money received can be used to cover the financial shortfall faced by Covid 19.

2.3. Knowledge, Attitude and Practices (KAP) of Zakat Recipients

The KAP survey is used to gather data on a certain population to determine their knowledge (K), attitudes (A), and practices (P) on a particular matter. Knowledge (K) as a collection of ideas, information, and "science." It's also a person's ability to imagine and see things. The survey's assessment of knowledge aids in identifying areas where further information and zakat activities are needed.

Meanwhile, attitude (A) indicates the variable that sits in the middle of the circumstance and the response to it. It explains why, out of all the possible responses to a stimulus, a subject chooses one. Besides, attitudes and behaviors are not readily observable, it is a good idea to examine them. It is worth noting that multiple studies have found little, if any, correlation between attitudes and behaviors. Furthermore, practices (P) also known as behaviors are the observable actions of an individual in response to a certain topic study. This is a topic that is concerned with the tangible, with activities. For example, in practices related to the zakat program, individuals gather information about challenges faced during the application process and the actions taken in response.

A KAP survey can produce information that can be useful to help develop and implement interventions, determine knowledge gaps, cultural attitudes, and behavioral patterns that may indicate needs, issues, and obstacles. It's also useful to gain a better understanding of commonly held beliefs, attitudes, and other factors that influence behavior. Lastly, it helps establish baselines and monitor changes resulting from interventions. The KAP assessment method was used by Oi & Norkiah (2011) to assess the level of knowledge, attitude and practices of food handlers in the cafeteria of residential colleges and canteens at Universiti Kebangsaan Malaysia on food hygiene and safety aspects. The results of such assessments show a positive attitude towards the items studied. Overall, out of all three levels, the survey respondents showed only good attitude in food handling and cafeteria.

Jerome et al. (2023) using the Knowledge, Attitude, and Practice (KAP) survey method gather information from the public on KAP in the context of Sarawak's growing popularity of indigenous popular music. Preliminary results showed that most of the respondents had sufficient knowledge of the

popular songs of the Ibans, Bidayuh, Melanau and Kenyah. Respondents showed a positive attitude towards the songs and were expressed through shared music and engagement choices such as listening, dancing, and singing. Respondents play a role played by the popular music of the indigenous people of Sarawak in the construction of identity. The results showed a significant correlation between knowledge, attitudes and practices with Sarawak's indigenous popular songs among Malaysians.

In studying the Covid-19 virus (Zur Batrisyia et al., 2021), a study was conducted among veterinary and medical students of the clinical year at a university in Malaysia to determine students' knowledge, attitude, and practice towards COVID-19. Overall, respondents showed high knowledge (80%), moderate attitude (76%) and high practice (86%) towards COVID-19. The analysis revealed a statistically significant association between attitudes and practices although there was no link between knowledge and attitudes and practice variables. Overall, the results showed that both veterinary and medical students had high knowledge with moderate to high attitudes towards COVID-19. Therefore, they adopt good preventive measures in limiting the spread of the disease.

Nurhasniza et al. (2022) applied the theory of Knowledge, Attitude, Practice (KAP) to assess public awareness of the COVID-19 pandemic with a direct impact on general practice, which influenced by its knowledge and attitude. The results from this study conclude that public knowledge related to COVID-19 prevention can influence how they develop certain attitudes and practices regarding COVID-19 prevention. Moreover, it is believed that people's awareness is important to create preventive trust, foster positive attitudes, and promote good practice. The existing KAP model is only used to understand and analyze human responses to certain phenomena, especially in the field of health studies but not considered the impact of social awareness on infectious diseases. Therefore, the original theory of KAP needs to be modified to include awareness in assessing the public's thinking on disease prevention measures.

While the Knowledge, Attitude and Practice (KAP) framework has been widely applied in health research, its application within Islamic financial practice particularly zakat management in higher education remains limited. Building on this gap, the present study adapts the KAP framework to an institutional zakat governance context by incorporating digital zakat systems and examining the relationships between KAP dimensions and student satisfaction. This context-specific application extends existing KAP usage by offering practical and policy-relevant insights rather than theoretical expansion.

Conceptually, the Knowledge, Attitude, and Practice (KAP) framework is not merely employed as a survey technique but as a behavioural model that explains how cognitive understanding influences evaluative dispositions and subsequently guides action. In this study, KAP functions as a conceptual lens to examine zakat engagement among higher education students, rather than solely as a descriptive measurement approach. Knowledge represents students' understanding of zakat eligibility, procedures, and institutional mechanisms; Attitude reflects evaluative beliefs toward zakat management, including perceptions of fairness, usefulness, and credibility; while Practice captures observable behaviours such as engagement with application processes and utilisation of zakat assistance. This conceptual positioning enables KAP to serve as an explanatory framework appropriate for analysing zakat-related behaviours within institutional contexts.

Therefore, the Knowledge, Attitude, and Practice (KAP) framework is not only applicable but also highly relevant for analyzing Islamic social finance mechanisms such as zakat, as it systematically explains how understanding of zakat principles and procedures translate into favorable attitudes and actual engagement with zakat systems among beneficiaries.

To further strengthen the theoretical grounding, the application of the Knowledge, Attitude, and Practice (KAP) framework in zakat management can be supported by established behavioral change and social psychology theories. In the context of zakat, knowledge (K) such as awareness of eligibility criteria, application procedures, and distribution mechanisms forms the cognitive foundation for individuals to

evaluate zakat as a legitimate and accessible form of assistance. This cognitive understanding directly influences attitude (A) by shaping perceptions of fairness, transparency, and effectiveness of zakat institutions, which in turn affects confidence and trust in the system.

This sequential relationship aligns with the Theory of Planned Behaviour (TPB), which posits that behaviour is driven by attitudes, subjective norms, and perceived behavioural control. In zakat management, students' positive attitudes toward zakat distribution, coupled with perceived ease of application (e.g., user-friendly e-zakat platforms), strengthen their intention to apply, ultimately resulting in actual practice (P) such as timely submission of applications and compliance with documentation requirements. Thus, KAP operationalizes TPB by empirically capturing the pathway from belief formation to action.

Similarly, Prochaska's Stages of Change Model provides further justification for KAP's applicability. Knowledge represents the pre-contemplation and contemplation stages, where students first become aware of zakat opportunities. Attitude corresponds with the preparation stage, reflecting motivational readiness and perceived benefits. Practice mirrors the action stage, where students actively engage in zakat application and utilization. This progression supports the use of KAP as a behavioral diagnostic tool to assess readiness, engagement, and sustainability of zakat usage among beneficiaries.

Within Islamic social finance, zakat is not merely a financial transfer but a faith-based social mechanism requiring conscious participation and trust. Therefore, the KAP framework is theoretically suitable for examining how religious knowledge and institutional communication shape beneficiaries' attitudes and behaviors. Applied to higher education settings, KAP enables zakat institutions to identify knowledge gaps, attitudinal constraints, and behavioral barriers, thereby enhancing governance efficiency, communication strategies, and equitable distribution outcomes.

Taken together, the reviewed studies highlight the fragmented examination of zakat-related behaviour, where knowledge, attitudes, practices, and satisfaction are often analysed separately. Addressing this gap, the present study adopts an integrated framework to examine how these dimensions interact from the perspective of zakat recipients in higher education institutions.

This study contributes at the institutional, measurement, and stakeholder levels by empirically validating university-based zakat management through an extended KAP –satisfaction framework, offering reliable measurement tools and actionable insights to strengthen inclusive education support for B40 students.



Figure 1. Conceptual Framework Knowledge, Practice in Zakat towards Satisfaction among B40 Students

Note: Satisfaction is conceptualized as an evaluative outcome of students' experiences with institutional zakat management, rather than as a component of attitude.

In the framework, Knowledge represents students' cognitive understanding of zakat eligibility, procedures, and institutional mechanisms. This knowledge informs Attitude, which reflects evaluative beliefs, motivation, and perceptions toward zakat management. These attitudinal dispositions subsequently shape Practice, manifested through behavioural engagement with the zakat application and utilisation process. Satisfaction is conceptualised as an evaluative outcome of these experiences, rather than as a component of attitude, and reflects students' overall assessment of institutional zakat management.

The conceptual framework assumes a directional sequence in which Knowledge forms the cognitive foundation influencing Attitude, which subsequently shapes Practice. Specifically, greater understanding of zakat procedures, eligibility criteria, and institutional mechanisms is expected to foster more favourable attitudes toward zakat management, leading to higher levels of engagement and

utilisation behaviour. While the present study examines associative relationships rather than causal effects, the proposed directionality is grounded in KAP theory and aligned with behavioural models such as the Theory of Planned Behaviour.

Although Attitude and Satisfaction are conceptually related, they are treated as analytically distinct constructs in this study. Attitude refers to students' evaluative beliefs and predispositions toward zakat as a support mechanism, encompassing perceptions of its role, relevance, and motivational value. In contrast, Satisfaction represents post-experience evaluations of institutional service delivery, including system accessibility, clarity of information, communication effectiveness, and administrative efficiency. By modelling Satisfaction as an outcome variable rather than a component of Attitude, this study avoids construct overlap and enables clearer interpretation of how cognitive and attitudinal factors translate into evaluations of zakat management performance.

The relationships proposed in the framework are theoretically grounded in behavioural models which posit that knowledge provides the cognitive basis for attitude formation, attitudes influence behavioural intentions and practices, and experienced practices subsequently shape evaluative judgments such as satisfaction. Within the context of zakat recipients, these relationships are further reinforced by beneficiaries' direct engagement with institutional procedures and service delivery mechanisms.

This study contributes to the zakat literature in three ways. First, it extends zakat research by focusing on beneficiaries rather than contributors, addressing an underexplored perspective within higher education settings. Second, it applies the KAP framework as an explanatory lens to understand behavioural engagement with zakat assistance. Third, by positioning Satisfaction as an outcome of institutional service delivery, the study offers practical insights for improving zakat management effectiveness among student recipients.

3. Methodology

3.1. Research Design

This study utilized the quantitative approach using a set of questionnaires encompassing KAP and satisfaction level inquiries to 236 B40 students at UiTM Negeri Sembilan to obtain data. The items of the set consist of the KAP of zakat recipients towards the zakat application system, Advertisement of zakat distribution, Sources of information, ZAWAF's programs and types of shocks by students. ZAWAF stands for Zakat, Sedekah, and Waqf Unit in UiTM. The detailed mapping of constructs, measurement items, and their sources is provided in Appendix to enhance transparency and replicability.

3.2. Sampling Method

In this study, zakat recipients refer to students classified under the B40 income group, as determined through institutional zakat eligibility assessments based on verified household income documentation and demonstrated financial need. As this study focuses on B40 zakat recipients, the findings may not be generalisable to students from higher income groups who do not qualify to receive zakat in UiTM Negeri Sembilan. This research focused on B40 students as the sample using a random stratified sampling method. This approach involves categorizing students based on their faculties at UiTM Negeri Sembilan and then randomly selecting participants from each stratum. By implementing this method, this study captured diverse perspectives from various faculties, enhancing the overall validity and generalizability of the findings. This meticulous sampling strategy ensures a well-rounded representation of B40 students across the university, contributing to a comprehensive understanding of their experiences and perceptions related to zakat management. This study was conducted on 236 students in the B40 category within UiTM Negeri Sembilan, which includes Kuala Pilah, Rembau and Seremban 3 campuses. The sample size of 236 students was gained from the records of the Hal Ehwal Pelajar (HEP) on the students who received zakat from the past four semesters.

3.3. Data Analysis

Descriptive analysis was applied to the demographic data of the samples, encompassing measures of frequency, central tendency, and dispersion. This approach aims to provide a comprehensive summary of the respondents' demographic characteristics within the sample size (Yellapu, 2018). Measures of frequency were employed to determine ratios and percentages of respondents answering specific questions, offering insights into response patterns (Yellapu, 2018). Additionally, the analysis yields mean, median, and mode values, providing a central tendency perspective within the sample (Yellapu, 2018). To assess the level of dispersion and identify the range of data, a measure of dispersion was utilized.

Quantitative analysis measured the Knowledge, Attitude, and Practice (KAP) levels, along with satisfaction levels among B40 students regarding zakat management at UiTM Negeri Sembilan. Cross-sectional questionnaires, employing a Likert-Scale method ranging from 1 = strongly disagree to 5 = strongly agree, was distributed. Reliability analysis using Cronbach's Alpha assessed the questionnaire's internal consistency, with a range of 0.7 to 0.8 considered acceptable (Ursachi et al., 2013). In measuring KAP levels, correlation analysis was conducted to determine the relationship between the KAP on zakat distribution by ZAWAF's. A statistical significance value ($p < 0.05$) was employed to assess the strength of the correlation.

The Likert-Scale method used to measure the level of satisfaction among B40 students, referencing the Wiersma minimum score (1995) for analysis (Zainal & Abu, 2019). The Statistical Package for Social Science (SPSS) was employed for the analysis, ensuring a robust and systematic evaluation of the collected data. This comprehensive data analysis approach aims to unveil insights into the demographic characteristics, KAP levels, and satisfaction levels among B40 students at UiTM Negeri Sembilan, contributing valuable findings for zakat management enhancement.

The analyses directly address the study objectives; overall KAP and Satisfaction levels, correlations among KAP variables and Satisfaction, and Satisfaction differences across campus, income, and CGPA were treated as confirmatory, while normality checks, reliability tests, and item-level mean patterns were considered exploratory.

In assessing Knowledge, Attitude, and Practice (KAP), ensuring measurement validity is essential to confirm that the survey items accurately capture the constructs intended. Like most self-reported KAP studies, this research is susceptible to common response biases such as social desirability bias, where students may provide favorable answers to appear compliant or appreciative of zakat assistance, and response-set tendencies (e.g., acquiescence or consistently choosing higher-end Likert options). To reduce these risks, the questionnaire was designed with clear, neutral wording and included both positively and negatively phrased items to encourage thoughtful responses. Participants were also assured of anonymity to minimize socially desirable responding. Nevertheless, these biases cannot be fully eliminated; therefore, they will be acknowledged as study limitations, and future research may incorporate mixed-methods triangulation or behavioral data to strengthen validity.

4. Results

This study aims to assess Knowledge, Attitude, and Practice (KAP) among B40 students regarding zakat management at UiTM Negeri Sembilan to get a quality education in their studies. Therefore, the results of this study are based on the research question as follows:

1. What are the levels of Knowledge, Attitude, Practice (KAP) toward zakat management among B40 students at UiTM Negeri Sembilan?
2. What are the differences in practice toward zakat management by campus, household income category and CGPA among B40 students at UiTM Negeri Sembilan?

3. What is the relationship between Knowledge, Attitude, and Practice (KAP) toward zakat management among B40 students at UiTM Negeri Sembilan?

4.1. Demographic of Respondents

Table 1. Summary of the Demographic of the Respondents

		Frequency (Number of Respondents)	Percent (%)
Gender	Male	58	24.6
	Female	178	75.4
Current Academic's Undertaking	Diploma	141	59.7
	Bachelor's Degree	95	40.3
Campus	Rembau	59	25.0
	Kuala Pilah	63	26.7
	Seremban 3	114	48.3
Household Income Classification	Less than RM2500	192	81.4
	RM2501 – RM3000	24	10.2
	RM3001 – RM3500	9	3.8
	RM3501 – RM4000	8	3.4
	RM4001 and above	3	1.3
CGPA	Less than 2.50	5	2.1
	2.51 – 3.00	52	22.0
	3.01 – 3.50	106	44.9
	3.51 and above	73	30.9
Number of Siblings	1	2	8
	2 – 3	73	30.9
	4 – 5	106	44.9
	6 – 7	38	16.1
	> 8	17	7.2

The summary of the demographic data of the respondents focuses on the respondent's background information. The variables of the demographic data consist of the respondent's gender, current academic undertaking, campus, household income classification, Cumulative Grade Point Average (CGPA) and number of siblings. Once the data collection process finished, the researcher managed to collect 236 respondents among the students of UiTM Negeri Sembilan.

According to the analysis, most of the respondents for this study is female with 75.4% (n=178) followed by male with 24.6% (n=58). Besides, the highest respondent of this study is the diploma students with 59.7% (n=141) while the remaining respondents are bachelor's degree students with 40.3% (n=95). The respondents from UiTM Seremban 3 are the highest contributor to this study with 48.3% (n=114) followed by UiTM Kuala Pilah with 26.7% (n=63) and UiTM Rembau with 25% (n=59). As for the household income classification, 81.4% (n=192) of the respondents answered that their household income is less than RM2,500 per month which equal to less than RM30,000 per annum. Then the data of household income classification followed by 10.2% (n=24) respondents came from family with income between RM2501 – RM3000, 3.8% (n=9) for respondents who came from family with income

between RM3001 – RM3500 and 3.4% (n=8) for the family with income between RM3500 – RM4000. Moreover, there are only 1.3% (n=3) of 236 respondents came from family with income above RM4001. Most of the respondents CGPA is between 3.01 to 3.50 with the percentage of 44.9% (n=106) followed by 3.51 and above with 30.9% (n= 73), 22% (n=52) of respondents for CGPA between 2.51 – 3.00 and 2.1% (n=5) with CGPA below 2.50. Finally, for the number of siblings there are 44.9% (n=106) of respondents are from family with the siblings of 4 -5, followed by the siblings of 2 – 3 with 30.9% (n=73), siblings of 6 – 7 with 16.1% (n=38) and 7.2% (n=17) for the respondents with the number of siblings more than eight. From the data collected there are only 8% (n=2) of the respondents who are the only child.

4.2. Normality Test

The normality test focused on skewness and kurtosis, which are measures of each variable's information. To test the normality of this data, the skewness value must be less than 2.0 (Hae, 2013).

Table 2. Test of Normality

Variables	Skewness	Remarks
Knowledge	-1.559	Normal
Attitude	-1.422	Normal
Practice	-1.911	Normal
Satisfaction	-1.644	Normal

The table above shows the test of normality for the three variables that were tested in this study. Among the four variables are knowledge, attitude, practice, and satisfaction. According to the table above, the result of knowledge is -1.559 for skewness. While for the attitude, the result shows that the skewness of this variable is -1.422. Besides, the skewness for the practice is -1.911, and the skewness for satisfaction is 1.644. Therefore, it can be concluded that the skewness results for all three variables are normal, since the skewness values are less than 2.0.

4.3. Reliability Test

Cronbach's alpha is the measure which is used by researchers in testing the reliability of the variables of the study. The range of accepted Cronbach alpha's value could be differed among the researchers where according to Nunnally 1967, it stated in his books where the value could be low as 0.50 and other researchers which is Hair et al (2016) prefer that 0.7 are the acceptable value for the value and could be as low as 0.60. According to George and Mallery (2003), there is a tiered approach in determining the acceptable value for Cronbach alpha which was used in this study. The tiered approach is as follows: > 0.9 (excellent) and ≤ 0.5 (unacceptable) (C. Howard, nd).

Table 3. Summary of Reliability Test

Variables	Cronbach's Alpha	No. of items
Knowledge	0.915	9
Attitude	0.868	7
Practice	0.936	8
Satisfaction	0.930	13

According to the table above, the value of Cronbach's alpha for knowledge is 0.915 with nine items or questions. The value of the Cronbach's alpha for knowledge is excellent, which is acceptable at 0.915. While for the attitude, the value of the Cronbach's alpha is also high, which is acceptable at 0.868, where

there are 7 items for the variables. Furthermore, the Cronbach's alpha for practice is also high at 0.936 (8 items). Lastly, the result for satisfaction shows that Cronbach's alpha value is in the acceptable range at 0.930 (13 items). In a nutshell, the tested variable for this study is shown as reliable where the value of Cronbach's alpha is within 0.868 and 0.930, where it's considered that the items or the questions created and answered are reliable.

4.4. The Level of Knowledge, Attitude, and Practice toward Zakat Management

The "Knowledge" section assesses the participants' awareness and understanding of zakat distribution processes, its importance, and its availability, particularly in a university setting. High mean scores in this section indicate a high level of zakat recipients' experience. If items related to knowledge regarding zakat distribution processes or the efforts made by the university in providing zakat assistance score highly, this implies that the university's communication and information dissemination about zakat are effective. Students are not only aware but also well-informed. This level of knowledge is crucial for ensuring that those who need zakat assistance can access it.

Table 4. Mean Analysis of Students' Knowledge

Items	Mean
B1) I am aware that there is an effort made by the university to help such as zakat distribution to the students.	4.64
B2) I am aware of the procedures to apply for zakat	4.49
B3) I am aware that B40 category students are entitled to receive a zakat.	4.74
B4) I know the duration of how long my zakat application could be approved.	3.90
B5) I am aware that my zakat application will undergo the process of eligibility by ZAWAF's for the HEP department.	4.53
B6) I am aware of the eligibility requirements to apply for zakat.	4.51
B7) I am aware that zakat applications are open for every early semester.	4.44
B8) I am aware that the application and any information about zakat could be accessed in the student portal.	4.52
B9) I am familiar with the documents that are required for a zakat application.	4.40

From the table, items *"I am aware that there is an effort made by the university to provide assistance such as zakat distribution to the students"* have a mean score of 4.64. In addition, the item *"I am aware that B40 category students are entitled to receive a zakat"* has a mean score of 4.74. This suggests that, on average, participants agreed or strongly agreed with these two statements. These following items have the score above 4.5; *"I am aware that my zakat application will undergo the process of eligibility by ZAWAF's for the HEP department"*, *"I am aware of the eligibility requirements to apply for zakat"*, *"I am aware that the application and any information about zakat could be accessed in the student portal"*.

Each section (Knowledge, Attitude, Practices) contains a list of items with corresponding mean scores. These mean scores represent the average response of the study participants to each item. The items are statements that participants likely rated on a scale. The "Attitudes" section measures the participants' beliefs, perceptions, and feelings towards zakat distribution. This section is crucial in understanding how students perceive the moral, social, and practical aspects of zakat.

Table 5. Mean Analysis of Students' Attitude

Items	Mean
C1) I believed that zakat distribution by ZAWAF could ease the burden among B40 students to pay the university fee.	4.61
C2) I believed that ZAWAF could ease the burden among B40 students to buy the learning essential.	4.64
C3) Throughout the semester, I will have no issues in meeting my everyday needs (such as food).	3.62
C4) I am pleased with the sum University provided by ZAWAF.	4.10
C5) After receiving the zakat from university from ZAWAF, I feel motivated to study.	4.49
C6) The student portal is easy to access for the zakat application process.	4.47
C7) The eligibility criteria to apply for zakat made the needy students easy to access for the application.	4.42

High mean scores on items that measure belief in the positive impact of zakat distribution, like easing financial burdens, indicate a positive perception and acceptance among students. It reflects a communal understanding of zakat as a beneficial and ethical practice within the student body.

Table 6. Mean Analysis on Students' Practices

Items	Mean
D1) I am committed to preparing the documents as required by ZAWAF	4.55
D2) I am committed to complying with the zakat application rules.	4.60
D3) I am aware of the invitation for the interview.	4.58
D4) I am always checking the application status.	4.50
D5) I agree with the total zakat amount given.	4.33
D6) I always follow the updates regarding the zakat application by ZAWAF.	4.36
D7) I won't delay using the zakat money for study.	4.54
D8) I am aware of the closing date of the application.	4.51

The "Practices" section evaluates the behaviors, actions, or practices of students in relation to zakat, such as their engagement in the application process or their commitment to fulfilling requirements for zakat distribution. High mean scores in practices, like commitment to preparing necessary documents, indicate that students are not only aware and have positive attitudes but are also actively engaging in the zakat process. It shows a practical application of their knowledge and attitudes.

The "Satisfaction" section assesses the participants' satisfaction with zakat management, covering information dissemination, system accessibility, processing duration, distribution methods, and overall institutional performance. High mean scores in this section indicate a high level of zakat management's performance. If items related to satisfaction with zakat management score highly, this implies that the zakat service is good and well delivered to the students. From the table, item "*E1) I feel comfortable using the e-zakat facility when applying for zakat*" has the highest mean score with a mean = 4.45. In addition, the item "The e-zakat system is user-friendly" has a mean score of 4.38. The other four items have mean value 4.30 and above are item "Method used by ZAWAF to distribute information about zakat application and zakat distribution is effective (e.g UiTM Student Portal, social media, WhatsApp, Telegram)", item "The instructions given in the e-zakat portal for zakat application are clear and easy", item "I feel comfortable with the preparation of documents during the application process through e-zakat" and item "The time given for students to apply for the zakat application with the method of information distributed is sufficient". This suggests that, on average, participants agreed or strongly agreed with e-zakat services and the method used by zakat management to distribute information about zakat.

Table 7. Mean Analysis on Students' Satisfaction

Items	Mean
E1) I feel comfortable using the e-zakat facility when applying for zakat.	4.45
E2) The e-zakat system is user-friendly.	4.38
E3) I feel comfortable with the preparation of documents during the application process through e-zakat.	4.35
E4) The dissemination of information regarding the zakat application on the university's social media is comprehensive.	4.17
E5) Information about the zakat application is quickly shared on the university's social media platforms.	4.14
E6) I discovered that university social media is slow in disseminating information about zakat applications.	3.31
E7) Social media makes it simple for me to learn more about the application.	4.18
E8) The information of zakat distribution by ZAWAF is well distributed.	4.20
E9) The instructions given in the e-zakat portal for zakat application are clear and easy.	4.36
E10) Method used by ZAWAF to distribute information about zakat application and zakat distribution is effective (e.g UiTM Student Portal, social media, WhatsApp, Telegram).	4.37
E11) The time given for students to apply for the zakat application with the method of information distributed is sufficient.	4.33
E12) The zakat application status given within the given period of time.	4.19
E13) The zakat sum given within the suitable period of time	3.94

4.5. Comparison of Mean Differences in Participation by Campus, Household Income and CGPA

Table 8. Mean Differences in Participation by Campus, Household Income and CGPA

Variable	n	Mean	SD	F	p
Campus					
Rembau	59	4.50	0.63	3.96	.02
Kuala Pilah	63	4.67	.423		
Seremban 3	114	4.39	.692		
Household Income				.314	.87
Less than RM2500	192	4.51	.59		
RM2501 – RM3000	24	4.43	.87		
RM3001 – RM3500	9	4.50	.57		
RM3501 – RM4000	8	4.39	.56		
RM4001 and above	3	4.20	.59		
CGPA				3.31	.02
Less than 2.50	5	3.70	1.53		
2.51 – 3.00	52	4.48	.51		
3.01 – 3.50	106	4.6	.56		
3.51 and above	73	4.46	.67		

A one-way analysis of variance was done to compare practices toward zakat management by three campuses of UiTM. The findings revealed the differences were significant, $F(236) = 3.96$, $p = .02$. Meanwhile, the findings showed that there is no difference in practice toward zakat management by household income group, with $F(236) = .314$, $p = .87$. The findings showed that there is significant difference in practice toward zakat management by CGPA, with $F(236) = 3.31$, $p = .02$.

4.6. The Relationship between Knowledge, Attitude, Practice and Satisfaction toward Zakat Management

A Pearson Product-Moment Correlation was performed to test the relationships between knowledge, attitude, practice and satisfaction as shown in Table 9. The results revealed a high, significant positive relationship ($r = 0.739$, $n = 236$, $p = .000$) knowledge and satisfaction toward zakat management, as well as a high, significant positive relationship between attitude and satisfaction toward zakat management ($r = .861$, $n = 236$, $p = .000$). In addition, the findings showed a high, significant positive relationship ($r = 0.681$, $n = 236$, $p = .000$) practice and satisfaction toward zakat management. Based on the findings, as the knowledge, attitude, and practice toward zakat management increase, the satisfaction of the respondents toward zakat management also increases.

Table 9. Correlation Coefficient Between Knowledge, Attitude, and Practice

Satisfaction	<i>r</i>	<i>p-value</i>
Knowledge	.739**	.000
Attitude	.861**	.000
Practice	.681**	.000

** Correlation is significant at the 0.01 level

5. Discussion

The findings should be interpreted in light of the respondents' socioeconomic position as B40 students, for whom financial constraints may influence knowledge, attitudes, practices, and satisfaction toward institutional zakat assistance.

This study assessed the levels of Knowledge, Attitude, and Practice (KAP) among B40 students regarding zakat management at UiTM Negeri Sembilan. The findings revealed that students possess a high level of awareness about the zakat application process, demonstrate positive attitudes towards its impact on their lives, and actively engage in the application procedures. This indicates that the Zakat management system implemented by ZAWAF is received by the students. This study also highlights the importance of clear communication, user-friendly application systems, and efficient dissemination of information through social media and student portals. Students' high satisfaction levels with zakat management show that zakat assistance not only alleviates financial burdens but also motivates them to excel academically.

The findings suggest that students' understanding of zakat eligibility and institutional procedures plays a foundational role in shaping their evaluative attitudes toward zakat management. Consistent with the conceptual framework, favourable attitudes toward zakat management were associated with more proactive engagement in zakat-related practices, including application compliance and utilization behaviour. The results further indicate that students' practical engagement with zakat processes significantly influenced their satisfaction with institutional zakat service delivery. Higher levels of satisfaction were associated with positive educational and psychosocial outcomes, such as improved academic continuity, reduced financial stress, and increased trust in zakat institutions.

The transfer of knowledge about zakat can lead to more efficient distribution and use of zakat funds. It also prepares students to be better informed about Islamic financial practices, which is particularly important in regions where Islamic finance plays a significant role. Positive attitudes towards zakat can foster a supportive environment for zakat practices within the university. It might also encourage students to participate in or contribute to zakat in the future, reinforcing a culture of mutual assistance and social responsibility. Zakat distribution needs to be carried out efficiently and reach the intended

recipients. It also reflects a proactive stance among students in managing their financial responsibilities and seeking available assistance.

A one-way ANOVA was conducted to determine whether practices toward zakat management differed across three UiTM campuses, household income groups, and CGPA categories. The results showed a statistically significant difference in zakat management practices among the three campuses with $F = 3.96, p = .02$. This indicates that students' practices toward zakat management differed across campuses. This situation revealed that institutional context may influence students' engagement with zakat-related practices. Differences in campus administration, exposure to zakat awareness programmes, accessibility of zakat units, communication channels, or local campus culture may explain these variations. This finding is consistent with a recent study by Bahori et al. (2024), who reported that students' knowledge, attitude, and practice toward zakat management in higher education are shaped by institutional delivery systems and the dissemination of zakat information through university zakat units. They highlighted that unequal access to information and services may create differences in student practices across campuses.

In contrast, the ANOVA results revealed no statistically significant difference in zakat management practices based on household income group, $F = .314, p = .87$. This suggests that students from different income backgrounds demonstrated relatively similar practices toward zakat management. One possible explanation is that zakat management behaviour among university students may be driven more strongly by religious obligation, awareness, and institutional support rather than socioeconomic background. Since zakat is fundamentally rooted in Islamic values and communal responsibility, students may hold comparable behavioural tendencies regardless of family income status. This finding aligns with the broader literature indicating that behavioural engagement in zakat-related matters is often influenced more by intrinsic values and perceived importance than by income level alone. Rosele et al. (2024) similarly emphasized that trust, awareness, and management efficiency are more decisive factors in zakat participation than purely economic status.

The findings also showed a statistically significant difference in zakat management practices according to CGPA, $F = 3.31, p = .02$. This suggests that academic performance is associated with different levels of zakat management practice. Students with higher CGPA may demonstrate stronger discipline, responsibility, and better information-processing skills, which can positively affect their financial management and compliance-related behaviours, including zakat practices. Academically stronger students may also be more responsive to institutional information and more aware of ethical or religious responsibilities. This supports the view that educational attainment and academic engagement can enhance responsible behavioural outcomes.

The Pearson Product-Moment Correlation analysis demonstrated that knowledge, attitude, and practice toward zakat management were all positively and significantly associated with respondents' satisfaction. Among the three variables, attitude showed the strongest relationship with satisfaction ($r = .861$), followed by knowledge ($r = .739$) and practice ($r = .681$). These findings indicate that respondents who possess better understanding of zakat management, hold more favorable perceptions toward zakat institutions, and engage in positive zakat-related practices tend to report higher satisfaction levels with zakat management services.

The strong positive relationship between knowledge and satisfaction suggests that when respondents are well informed about zakat procedures, distribution mechanisms, eligibility criteria, and institutional roles, they are more likely to appreciate and trust the management system. This finding aligns with recent studies emphasizing that public awareness and understanding are essential in strengthening confidence toward zakat institutions. For example, Haslina et al. (2024) noted that effective management of zakat institutions requires greater stakeholder confidence, which can be enhanced through transparency and public understanding of zakat operations. Likewise, Rosele et al. (2024) reported that better public awareness of zakat systems contributes to stronger trust and confidence among zakat payers.

The strongest correlation found between attitude and satisfaction indicates that respondents' perceptions, beliefs, and emotional evaluations toward zakat institutions play a central role in shaping satisfaction. Individuals who perceive zakat institutions as trustworthy, fair, efficient, and socially responsible are more likely to feel satisfied with the services provided. This result is consistent with Zulfikri et al. (2023), who found that satisfaction with zakat distribution, service quality, and disclosure practices are major determinants influencing trust in zakat institutions. Positive attitudes often emerge when institutions demonstrate accountability and professionalism, leading to greater satisfaction among stakeholders.

The significant relationship between practice and satisfaction also implies that respondents who actively engage in zakat-related behaviors such as timely payment, use of institutional payment channels, and participation in zakat programs tend to experience higher satisfaction. Practical interaction with zakat systems may expose users to efficient services, digital platforms, and convenient payment methods, thereby enhancing satisfaction. This is supported by Haslina et al. (2025), who found that quality-in-use dimensions of online zakat systems significantly affect user satisfaction, highlighting the importance of user-friendly digital platforms in zakat management. Similarly, Olivia et al. (2024) concluded that digital technology in zakat management can improve trust and loyalty among contributors.

Overall, these findings support the Knowledge-Attitude-Practice (KAP) framework, where knowledge shapes favorable attitudes, which subsequently influence practices and satisfaction. The stronger effect of attitude suggests that while knowledge and practice are important, emotional trust and positive perception toward zakat institutions are especially crucial. Therefore, zakat institutions should prioritize educational campaigns, transparent governance, responsive service delivery, and effective communication strategies to improve public perceptions and satisfaction.

From a practical perspective, institutions managing zakat should focus on enhancing public literacy regarding zakat management processes, increasing service efficiency, and strengthening institutional integrity. When respondents perceive zakat institutions as competent and trustworthy, their satisfaction is likely to increase, which may also encourage continued zakat compliance and stronger support for institutionalized zakat systems.

The findings of this study can contribute to the decision of policy and communication. The study's findings can guide the ZAWAF's to enhance the university policies and to develop more communication strategies regarding zakat distribution, ensuring they are aligned with students' knowledge levels, attitudes, and practices. Furthermore, the finding not just contributes to the university's policies but shows reflection of broader cultural values as perceived by the students such as follows:

5.1. Reflection of Social Responsibility at Universities

High levels of knowledge and positive attitudes towards zakat among students might reflect broader cultural values emphasizing charity, social responsibility, and communal support in UiTM Negeri Sembilan. This could indicate a deeply ingrained cultural norm that helps those in need, especially through structured means like zakat, is highly valued (Abdillah et al., 2024). In UiTM, ZAWAF plays a significant role, perhaps the understanding of how students view and practice zakat can offer insights into how religious teachings influence contemporary social practices. It shows religious influence on social practices or the integration of religious obligations with everyday life, highlighting the role of faith in shaping social behaviors and norms.

5.2. Social Dynamics within the University

Social Cohesion and Mutual Support: The attitudes and practices surrounding zakat distribution can foster a sense of community and mutual support within the university. If students hold positive views and actively engage in zakat practices, it could lead to a more cohesive student body that values mutual

assistance. Awareness of Socio-economic Diversity: Students' engagement with zakat could also reflect their awareness and sensitivity towards socio-economic diversity and inequality among their peers. This understanding can foster a more inclusive environment where students are more attuned to the challenges faced by others and are willing to support them. It is same as Mat Daud & Wahid (2024) findings found that student readiness in the education process plays a significant role in mediating the effect between zakat aid and the learning process and academic performance.

5.3. Promoting Culture and Social Impact

Preparing Future Contributors and Beneficiaries: University students are future leaders and active members of society. Their perspectives on zakat can shape how they contribute to and engage with social welfare schemes in their professional and personal lives. Positive engagement at the university level could lead to continued support and participation in similar practices beyond the university, which in turn leads to social change and continuity (Yanmei et al, 2023). Social Change and Continuity: Students' attitudes towards zakat can be a bellwether for broader social change. If they hold progressive or differing views from traditional interpretations of zakat, this could signal evolving social norms. Conversely, strong adherence to traditional views can signify a desire for continuity in cultural and religious practices.

5.4 Educational and Policy Implications

Role of Education in Shaping Social Values: The way zakat is taught and discussed in educational settings can significantly influence students' understanding and attitudes. This highlights the role of education in shaping social values and practices related to charity and social welfare (Siti Aisyah et al., 2024). Informing Policy Development: Understanding students' KAP towards zakat can guide policymakers in developing relevant social welfare policies and programs that are culturally sensitive in addressing the needs of different community segments.

Thus, examining students' KAP towards zakat distribution not only provides insights into their individual beliefs and behaviors but also offers a window into the larger cultural and social fabric of the university community and, by extension, the wider society. It highlights how religious and cultural values are manifested in contemporary social practices and attitudes, and how these might evolve with future generations. In part of educational opportunities, these findings could also highlight areas where more education or information is needed, guiding future programs or interventions to improve students' understanding and engagement with zakat. All findings should be interpreted within the context of B40 student's experiences as zakat recipients' as the respondents were beneficiaries of zakat assistance rather than representatives of the general student population.

6. Conclusion

In conclusion, this study provides a deeper understanding of how B40 students at UiTM Negeri Sembilan interact with the zakat system provided. The findings can serve as a foundation for improving ZAWAF's policies and communication strategies to be more inclusive and responsive to students' needs. Moreover, the study reflects the strong cultural and religious values among students, particularly in terms of social responsibility and community support. Therefore, efficient zakat management not only supports B40 students financially but also nurtures a generation that is socially conscious and committed to helping others. These conclusions reflect the experiences of university B40 students who received zakat assistance and should be understood within the scope of beneficiary-based engagement rather than general student perceptions. This study demonstrates that zakat engagement among B40 university students is best understood as a behavioural process that begins with knowledge, progresses through attitudinal evaluation and practical engagement, and culminates in satisfaction with institutional zakat management. Satisfaction subsequently contributes to broader intended outcomes, including academic continuity and enhanced trust in zakat institutions.

Acknowledgement

This article funded by research grant entitled *Knowledge, Attitude and Practice (KAP) and Satisfaction Survey on Zakat Management in UiTM Negeri Sembilan* project reference code 100-UiTMKNS (ZAWAF (GPZ. 02/2023) as well as Academy Contemporary of Islamic Studies, Universiti Teknologi MARA (UiTM) Cawangan Negeri Sembilan.

Ethics Statement

This study has received permission and approval from the Research Management Centre, UiTM Negeri Sembilan. All personal data of respondents is confidential.

References

- Abdillah, A., Kassim, Z. A., Drahman, D. N. A., Hossen, R., & Yon, W. A. W. (2024). Management of Zakat Funds by Students in a Public University. *International Journal of Academic Research in Business and Social Sciences*, 14(11), 484–493. <http://dx.doi.org/10.6007/IJARBS/v14-i11/23181>
- Abdul, H. M. N., Mohammed, S. A. R., Rashidah, A. R., Rozman, M. Y., & Siti, M. A. (2012). *Assessing performance organizations: A framework for Zakat Institutions*. *British Journal of Economics, Finance and Management Sciences*, 5(1), 12-22.
- Alam, G. M. (2023). *Sustainable Education and Sustainability in Education: The Reality in the Era of Internationalisation and Commodification in Education—Is Higher Education Different?* *Sustainability*, 15(2), 1315. <https://doi.org/10.3390/su15021315>
- Allang, B.A., Awang, M.M., Ahmad, A.R., & Ahmad, A. (2019). *Influenced Factors of B40 Students' Academic Achievement*. Proceedings of The 2nd International Conference on Sustainable Development & Multi-Ethnic Society - ICOSH2 2019.
- Anwar, M., & Roziyah, S. (2021). *Komitmen Majlis Agama Islam Negeri Sembilan (MAINS) dalam era covid 19 di Negeri Sembilan*. *Journal of Fatwa Management and Research*, 25(1), 34-48.
- Aziz, S. N. A., Mohd Zahid, E. S., Mearaj, M., Sumery, Z., Redzuan, M. A., & Abdul Kadir, J. (2020). Tadbir Urus Agihan Dana Zakat di UiTM Cawangan Johor: Kajian Terhadap Isu dan Penambahbaikan Pengurusan Agihan. *Journal of Islamic Philanthropy & Social Finance (JIPSF)*, 1(1), 9-22.
- Azli, R. M., Kamaruzzaman, M. A. S., & Zhaffar, N. M. (2024). *Analisis ekonomi pelajar: Tinjauan terhadap permohonan zakat di UiTM Negeri Sembilan*. *Journal Islamic Philanthropy and Social Finance*, 6(1), 94-107.
- Azman, A. R., & Ahmad, W. M. (2011). *Kelayakan dan kedudukan pelajar sebagai penerima zakat menurut perspektif Islam*. *Jurnal Pengurusan JAWHAR, Jabatan Wakaf, Zakat & Haji, Jabatan Perdana Menteri*, 5(1), 131-147.
- Azman, A. R. (2012). *Zakat institutions in Malaysia: Problem and issues*. *Global Journal Al-Thaqafah*, 2(1), 35-41.
- Bahori, A., Mohd Azli, R., Yaakob, A. F., & Mohd Yunos, Z. (2024). *Knowledge, Attitude and Practice (KAP) on Zakat Management in Higher Education*. *Environment-Behaviour Proceedings Journal*, 9(SI22), 51–56. <https://doi.org/10.21834/e-bpj.v9iSI22.5800>
- Hae, K. (2013). *Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis*. *Restoring Dentistry & Endodontics*, 52-54. <http://doi:10.5395/rde.2013.38.1.52>
- Hamzah, N., Masroom, M. N., Ruskam, A., & Saim, A. I. (2021). Penggunaan Wang Bantuan Zakat dalam Kalangan Asnaf Fisabilillah dan Asnaf Gharimin di Universiti Teknologi Malaysia. *Azka International Journal of Zakat & Social Science*, 1(2), 187-209.
- Hainnuraqma, R., Hamidah, M. I., Nor Balkish, Z., Mohamed Fairouz, A.K., & Sobhani, F.A. (2024). Perceived usefulness and attitude toward intention and acceptance of e-payment zakat. *Journal of Nusantara Studies*, 9(1), 320-344. <http://dx.doi.org/10.24200/jonus.vol9iss1pp320-344>
- Haslina Mohd, Baharom, F., Mohamed, S. F. P., Yusuf, Y., Abai, N. H. Z., Barraod, S. O., & Ba-Quttayyan, B. (2025). User satisfaction of the zakat system in the Kedah state: From a quality-in-use perspective. *Journal of Digital System Development*, 3(1). <https://doi.org/10.32890/jdsd2025.3.1.10>

Howard, C. (n.d.). Introduction to Cronbach's Alpha. Retrieved from mattchoward.com: [https://mattchoward.com/introduction-to-cronbachs-alpha/]

Ibrahim, M. F. (2015). *Sistem pengurusan zakat di Malaysia: Analisis strategi penyebaran maklumat menerusi laman sesawang [Zakat management systems in Malaysia: Analysis of information dissemination strategy through website]*. Jurnal Pengurusan, 42, 119-130.

Izani, M., & Aripin, M. A. (2017). *Proses agihan zakat secara efektif di Universiti Awam Malaysia: Pengalaman UiTM Cawangan Sabah [The effective zakat distribution process at the public university of Malaysia: UiTM Sabah branch experience] [Paper presentation]*. Konferensi Antarabangsa Islam Borneo KAIB, Balik Papan Samarinda Indonesia, 2017, September 25.

Jerome, C., Ting, S. H., Perry, E. J., & Lim, C. K. N. (2023). *Knowledge, attitudes, and practices (KAP) towards popular indigenous music of Sarawak: A preliminary survey*. International Journal of Social Science Research, 5(2), 17-30.

Mat Daud, M. S., & Wahid, H. (2024). Assessment of educational performances: the role of zakat and learning-teaching process in higher education institution in Malaysia. Journal of Islamic Accounting and Business Research, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/JIABR-04-2024-0146>

Muda, Khadijah and Awang, Mohd Daud and Abdul Rahim @ Abdul Wahab, Husniyah and Abang Abai, Dayang Shobihah (2013) Persepsi pengguna laman web LZS terhadap pusat Lembaga Zakat Selangor dan pembayaran zakat melalui atas talian. In: Seminar Hasil Penyelidikan Sektor Pengajian Tinggi ke-3 2013, 2-3 Jul. 2013, EDC, Universiti Utara Malaysia. (pp. 1792-1803).

Muhammad Tajuddin Abdul Rahim & Wan Mohd Al Faizee Wan Ab Rahaman. (2017). *Pengurusan zakat IPT: Pengalaman Unit Zakat UKM*. [https://www.researchgate.net/publication/321125194_Pengurusan_Zakat_IPT_Pengalaman_Unit_Zakat_UKM]

Mustakim, Siti Salina and Hassan, Aminuddin and Chen, Xiaofei (2024). *Economic environment in fostering the internationalization of higher education institutions: what matters?* International Journal of Academic Research in Business and Social Sciences, 14 (2). pp. 332-342. ISSN 2222-6990; eISSN: 2308-3816

Nor Samsiah Sani, Ahmad Fikri Mohamed Nafuri, Zulaiha Ali Othman, Mohd Zakree Ahmad Nazri and Khairul Nadiyah Mohamad. (2020). *Drop-Out Prediction in Higher Education Among B40 Students*. International Journal of Advanced Computer Science and Applications (IJACSA), 11(11), 550-559. <http://dx.doi.org/10.14569/IJACSA.2020.0111169>

Nurhasniza Idham Abu Hasan, Shahrman Zainal Abidin, Mogana Darshini Ganggayah, Nur Faezah Jamal, & Wan Noor Hayatie Wan Abdul Aziz. (2022). *Knowledge, attitude and practices (KAP) theory towards preventive measures among Malaysians in early outbreak of COVID-19*. Malaysian Journal of Public Health Medicine, 22(1), 38-47.

Olivia, H., Hayati, I., Afandi, A., Daud, A., & Pasha, F. A. (2024). Integration of digital technology in zakat management: From accountability to muzzaki loyalty through trust. *Journal of Law and Governance*, 7(1).

Omar, P. M. F. F. A., Hajimin, M. N. H. H., Lada, S., Ghazali, H. M., & Khalil, N. M. (2022). *The acceptance of online zakat payment among Muslims in East Malaysia: A conceptual study*. Labuan E-Journal of Muamalat and Society, 22, 13-19.

- Paizin, M. N., & Som, H. M. (2019, September 24). *Penerimaan masyarakat untuk membayar zakat menerusi saluran atas talian (Online): Satu tinjauan baharu [Community receiver to pay zakat through online channels (Online): A new review] [Paper presentation]*. 5th Muzakarah Fiqh & International Fiqh Conference, MFIFC Kolej Universiti Islam Antarabangsa Selangor.
- Rosele, M. I., Muneem, A., Che Seman, A., Abdullah, L. H., Abdul Rahman, N. N., Abd Sukor, M. E., & Ali, A. K. (2024). The status quo of zakat management in Malaysia: Obstacles, opportunities and directions. *Labuan E-Journal of Muamalat and Society*, 18(1), 94–110.
- Shaharin, N. S., Bhari, A., Mohamed Yusof, M. F., & Yaakob, M. A. Z. (2021). Analisis Agihan Zakat di Lembaga Zakat Selangor Semasa Pandemik Covid-19. Perdana: International Journal of Academic Research (Social Sciences & Humanities), 11(1), 1-10.
- Shamsuddin, K. B., Wahab, B., & Salleh, M. S. (2021). *Impak bantuan zakat terhadap prestasi akademik pelajar dan penghayatan agama: Kajian kes semasa era Covid-19*. Proceedings of the 2nd Kedah International Zakat Conference 2021 (i-KEIZAC 2021) eISBN: 978967171284841. (Impact of zakat assistance on the students' academic performance and religiosity: A case study during the Covid-19 era)
- Sharif, D., Rameli, M. F. P., Aris, N. M., Adnan, S. D. M., Nasir, M. F. M., & Abas, S. A. (2022). *Perceptions on e-Zakat advancement distribution concept among UiTM Melaka staff: A qualitative study*. International Journal of Academic Research in Economics and Management Sciences, 11(3), 698-709.
- Siti Aisyah Zahari & Shahida Shahimi & Suhaili Alma'amun & Mohd Mursyid Arshad. (2024). "Millennials and Gen-Z Ethical Banking Behaviour in Malaysia," Springer Books, in: Zulkefly Abdul Karim & Ruzita Abdul Rahim & Wai Yan Wong & Siti Farah Dilla Zakaria (ed.), Contemporary Issues in Finance, Investment and Banking in Malaysia, pages 161-177, Springer.
- Takril, N.F, & Othman, N. F. (2020). *Effectiveness Of Zakat Collection and Distribution of Zakat Counter at Higher Learning Institutions*. International Journal of Islamic Economics and Finance Research, 3(2), 91–106. <https://doi.org/10.53840/ijiefer35>
- Ursachi, G., Horodnic, I. A., & Zait, A. (2013). *How reliable are measurement scales? External factors with indirect*. Procedia Economics and Finance, 20, 679-686.
- Yanmei, Z., Muhamad, M. M., Mustakim, S. S., & Zhufen, Y. (2023). *Strategies for Developing Holistic University Students within the Framework of Five-dimensional Education Integration in China: Status, Concept, and Realization Paths*. International Journal of Academic Research in Progressive Education and Development, 12(2), 426–439.
- Yellapu, V. (2018). *Descriptive statistics*. International Journal of Academic Medicine, 4(1), 60-63.
- Zainal, H., & Abu, B. A. (2019). *Tahap kepuasan agihan dalam kalangan pihak berkepentingan zakat di Negeri Kedah*. International Journal of Zakat and Islamic Philanthropy, 1(1), 14-19.
- Zulfikri, Kassim, S., & Othman, A. H. A. (2023). Exploring zakat payers' attitudes determinants on trust in zakat institutions. *Journal of Islamic Social Finance*, 1(2).
- Zur Batrisyia Mohd Zubir, Nor Yasmin Abd Rahaman, Rukman Awang Hamat, & Hayati Kadir @ Shahrar. (2021). *University students' knowledge, attitude, and practice (KAP) towards COVID-19 in Malaysia*. PERTANIKA Journal of Social Science and Humanities, 29 (4). 2853 - 2868. ISSN 0128-7702; ESSN: 2231-8534.

APPENDIX

SECTION A: DEMOGRAPHIC OF RESPONDENTS

SEKSYEN A: DEMOGRAFI RESPONDEN

1. Gender / *Jantina*

Male / *Lelaki* ☐ Female / *Perempuan* ☐

2. Current Academic's Undertaking / *Tahap Pendidikan*

Diploma / *Diploma* ☐ Degree / *Ijazah Sarjana Muda* ☐

3. Campus / *Kampus*

Rembau ☐ *Kuala Pilah* ☐ *Seremban 3* ☐

4. Household Income Classification / *Klasifikasi Pendapatan Isirumah*

Less than RM 2500 / *Kurang dari RM2500* ☐

RM2501 - RM3170 ☐

RM3171 – RM3970 ☐

RM3971 – RM4850 ☐

No Income / *Tiada pendapatan* ☐

5. CGPA / *Keputusan Peperiksaan Kumulatif*

Less than 2.50 / *Kurang daripada 2.50* ☐

2.51 – 2.99 ☐

3.00 – 3.50 ☐

3.51 and above / *3.51 dan keatas* ☐

6. Number of siblings / *Jumlah Bilangan Adik Beradik*

1 ☐

2 – 3 ☐

4 -5 ☐

6 – 7 ☐

> 8 ☐

SECTION B: KNOWLEDGE**SEKSYEN B: PENGETAHUAN**

1	2	3	4	5
Strongly Disagree / Sangat Tidak Setuju	Disagree / Tidak Setuju	Mixed Feeling / Tidak Pasti	Agree / Setuju	Strongly Agree / Sangat Setuju

No.	Questions / Soalan
1	I am aware that there are effort made by university to provide assistance such as zakat distribution to the students. / <i>Saya mengetahui bahawa terdapat usaha yang dilakukan oleh pihak universiti untuk menyediakan bantuan seperti pengagihan zakat kepada pelajar.</i> (Nor Muhamad & Abu Hassan, 2022)
2	I am aware of the procedures to apply for zakat / <i>Saya mengetahui tentang prosedur untuk memohon zakat.</i> (Ab Rahman, Mokhtar, Wan Razali, Ahmad, & Syed Omar, nd)
3	I am aware that B40 category students are entitled to receive a zakat. / <i>Saya mengetahui bahawa pelajar kategori B40 layak untuk mendapat zakat.</i> (Ab Rahman, Mokhtar, Wan Razali, Ahmad, & Syed Omar, nd)
4	I know the duration of how long my zakat application could be approved / <i>Saya mengetahui tempoh masa permohonan zakat saya akan diluluskan.</i> (Ab Rahman, Mokhtar, Wan Razali, Ahmad, & Syed Omar, nd)
5	I am aware that my zakat application will undergo the process of eligibility by ZAWAF or HEP department. / <i>Saya mengetahui bahawa permohonan zakat saya akan melalui proses semakan, tapisan dan kelayakan oleh pihak ZAWAF ataupun pihak HEP.</i>
6	I am aware of the eligibility requirements to apply for zakat / <i>Saya mengetahui tentang syarat kelayakan untuk memohon zakat.</i>
7	I am aware that zakat applications are open for every early semester / <i>Saya mengetahui bahawa permohonan zakat dibuka pada setiap awal semester.</i>

8	I am aware that the application and any information about zakat could be accessed in the student portal. / <i>Saya mengetahui bahawa permohonan dan segala informasi tentang zakat boleh didapati di portal pelajar.</i>
9	I am familiar with the documents that required for zakat application / <i>Saya mengetahui mengenai dokumen-dokumen yang diperlukan untuk memohon zakat.</i>

SECTION C: ATTITUDE (REVIEW / RECONSTRUCT)**SEKSYEN C: SIKAP (RELEVENCY FOR ATTITUDE QUESTIONS)**

1	2	3	4	5
Strongly Disagree / Sangat Tidak Setuju	Disagree / Tidak Setuju	Mixed Feeling / Tidak Pasti	Agree / Setuju	Strongly Agree / Sangat Setuju

No.	Questions / Soalan
1	I believed that zakat distribution by ZAWAF could ease the burden among B40 students to buy the learning essential / <i>Saya percaya bahawa agihan zakat dapat meringankan beban pelajar B40 untuk membeli barang keperluan pelajaran.</i>
2	Students that eligible to received zakat is the students who are from the B40 household income group / <i>Pelajar yang layak untuk menerima zakat adalah pelajar daripada golongan isirumah B40.</i>
3	Students who received zakat distribution will feel ease in fulfilling the daily essential throughout the semester (e.g food) / <i>Pelajar yang menerima zakat akan merasa beban berkurang dalam memenuhi keperluan harian sepanjang semester (contohnya makanan)</i>
4	Other than buying the essentials, students who received zakat distribution feel motivated to study. / <i>Selain daripada membeli barang keperluan, pelajar yang menerima zakat akan bertambah motivasi untuk belajar.</i>
5	Student's who are not comes from B40 category but faced with hardship to survive financially during semester are also could apply for zakat with valid reason / <i>Pelajar yang bukan daripada kalangan kategori B40 tetapi menghadapi kesukaran kerana masalah kewangan juga boleh memohon zakat dengan sebab yang kukuh.</i>

SECTION D: PRACTICE**SEKSYEN D: PRAKTIS**

1	2	3	4	5
Strongly Disagree / Sangat Tidak Setuju	Disagree / Tidak Setuju	Mixed Feeling / Tidak Pasti	Agree / Setuju	Strongly Agree / Sangat Setuju

No.	Questions / Soalan
1	I received the zakat application status and distribution process within the given period of time / <i>Saya menerima status permohonan dan proses pengagihan dalam masa yang telah ditetapkan.</i>
2	The information of zakat distribution by ZAWAF is well distributed / <i>Saya berpendapat bahawa informasi tentang pengagihan zakat oleh ZAWAF telah disampaikan dengan baik.</i> (Toto Ngadiman, Wahid, & Mohd Nor, 2016)
3	The instruction given in e-zakat portal for zakat application is clear and easy / <i>Saya berpendapat bahawa arahan yang diberikan di dalam portal e-zakat untuk memohon zakat adalah jelas dan mudah.</i> (Saud & Abd. Khafidz, 2018)
4	Method used by ZAWAF to distribute the information about zakat application and zakat distribution is effective (e.g UiTM Student Portal, Social Media, WhatsApp, Telegram) / <i>Kaedah yang digunakan oleh ZAWAF untuk menyampaikan informasi berkaitan permohonan dan pengagihan zakat adalah berkesan.</i>
5	The student portal is easy to be access for the zakat application process / <i>Portal pelajar adalah mudah untuk diakses bagi tujuan proses permohonan zakat.</i>
6	The eligibility criteria to apply for zakat made the needy students easy to access for the application / <i>Kriteria kelayakan untuk memohon zakat membuatkan pelajar yang memerlukannya mudah untuk mengakses permohonan zakat.</i>
7	The time given for students to apply for the zakat application with the method of information distributed is sufficient / <i>Masa yang diberikan kepada pelajar untuk membuat permohonan dengan informasi yang telah disampaikan adalah mencukupi.</i>

SECTION E: LEVEL OF STUDENT' S SATISFACTION**SEKSYEN E: TAHAP KEPUASAN PELAJAR**

1	2	3	4	5
Strongly Disagree / Sangat Tidak Setuju	Disagree / Tidak Setuju	Mixed Feeling / Tidak Pasti	Agree / Setuju	Strongly Agree / Sangat Setuju

No.	Questions / Soalan
1	I am satisfied with the method used by ZAWAF to distribute the information related to zakat application. / <i>Saya berpuas hati dengan kaedah yang digunakan oleh ZAWAF dalam menyampaikan maklumat mengenai permohonan zakat.</i>
2	I am satisfied with the system used in the Zakat application process. / <i>Saya berpuas hati dengan sistem yang digunakan di dalam proses permohonan zakat.</i> (Toto Ngadiman, Wahid, & Mohd Nor, 2016)
3	I am satisfied with ZAWAF initiatives taken by ZAWAF in distributing the information to the students. / <i>Saya berpuas hati dengan usaha yang diambil oleh ZAWAF dalam menyalurkan maklumat kepada pelajar.</i>
4	I agree with the eligibility criteria for the students to apply for Zakat. / <i>Saya bersetuju dengan syarat-syarat kelayakan pelajar untuk memohon zakat.</i>
5	I am satisfied with the accessibility to access the student portal for the zakat application process / <i>Saya berpuas hati dengan kebolehcapaian dalam mengakses portal pelajar untuk memohon zakat.</i>
6	I am satisfied with the duration of the zakat application process / <i>Saya berpuas hati dengan tempoh pemprosesan zakat.</i>
7	I am satisfied with the method used by ZAWAF to distribute the zakat fund to the students. / <i>Saya berpuas hati dengan kaedah yang digunakan ZAWAF dalam mengagihkan dana zakat kepada pelajar</i>
8	I am satisfied with the clear manual provided by ZAWAF to guide me to make the application in the student portal. / <i>Saya berpuas hati dengan manual yang jelas yang telah disediakan oleh ZAWAF sebagai panduan pelajar untuk memohon zakat.</i>

9	Overall, I am satisfied with zakat management in UiTM Negeri Sembilan. / Secara keseluruhan, saya berpuas hati dengan pengurusan zakat di UiTM Negeri Sembilan.

Mapping of Survey Items to Constructs

Construct	Item Code	Survey Item (Short Description)
Knowledge	B1	Awareness of university efforts to provide zakat assistance
	B2	Awareness of procedures to apply for zakat
	B3	Awareness that B40 students are entitled to receive zakat
	B4	Knowledge of expected duration for zakat approval
	B5	Awareness that applications undergo eligibility screening by ZAWAF/HEP
	B6	Awareness of eligibility requirements for zakat
	B7	Awareness that zakat applications open every early semester
	B8	Awareness that zakat information is accessible via the student portal
	B9	Familiarity with documents required for zakat application
Attitude	C1	Belief that zakat eases tuition-fee burden
	C2	Belief that zakat eases burden of purchasing learning essentials
	C3	Belief that zakat eases daily living needs during the semester
	C4	Satisfaction with the amount of zakat provided
	C5	Motivation to study after receiving zakat
	C6	Ease of accessing the student portal for zakat application

	C7	Perceived accessibility of eligibility criteria
Practice	D1	Commitment to preparing required documents
	D2	Compliance with zakat application rules
	D3	Awareness of interview invitations
	D4	Actively checking zakat application status
	D5	Agreement with total zakat amount received
	D6	Following updates regarding zakat applications
	D7	Prompt use of zakat funds for study purposes
	D8	Awareness of zakat application closing date
Satisfaction	E1	I feel comfortable using the e-zakat facility when applying for zakat.
	E2	The e-zakat system is user-friendly.
	E3	I feel comfortable with the preparation of documents during the application process through e-zakat.
	E4	The dissemination of information regarding the zakat application on the university's social media is comprehensive.
	E5	Information about the zakat application is quickly shared on the university's social media platforms.
	E6	I discovered that university social media is slow in disseminating information about zakat applications.
	E7	Social media makes it simple for me to learn more about the application.
	E8	The information of zakat distribution by ZAWAF is well distributed.
	E9	The instructions given in the e-zakat portal for zakat application are clear and easy.

	E10	Method used by ZAWAF to distribute information about zakat application and zakat distribution is effective (e.g UiTM Student Portal, social media, WhatsApp, Telegram).
	E11	The time given for students to apply for the zakat application with the method of information distributed is sufficient.
	E12	The zakat application status given within the given period of time.
	E13	The zakat sum given within the suitable period of time

Construct Operationalisation

In this study, the survey instrument operationalised four main constructs: Knowledge, Attitude, Practice (KAP), and Satisfaction with zakat management among B40 students at UiTM Negeri Sembilan. The Knowledge construct was measured using nine items (B1–B9) assessing students’ awareness and understanding of zakat availability, eligibility, procedures, timelines, and access to information via institutional platforms. Attitude was operationalised through twenty items (C1–C20) capturing students’ beliefs, perceptions, comfort levels, and evaluative judgments regarding the effectiveness, fairness, usability, and communication of zakat management and the e-zakat system. The Practice construct comprised eight items (D1–D8) reflecting students’ actual behaviours and engagement in the zakat process, including compliance with requirements, information-seeking behaviours, and utilisation of zakat funds. Finally, Satisfaction was measured using nine items (E1–E9) evaluating students’ overall satisfaction with zakat management, covering information dissemination, system accessibility, processing duration, distribution methods, and overall institutional performance. Collectively, these constructs align with the KAP framework, enabling systematic assessment of students’ awareness, perceptions, behaviours, and evaluative outcomes related to zakat management in a higher education context.

Factors Influencing Acceptance of CEFR-Aligned English Language Curriculum at a Malaysian University

ABSTRACT

Authors: Muhammad Fazli Abdullah^{1,2}, Fazilah Razali^{1*} and Farah Aida Fadzil^{1,2}

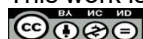
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

²School of Liberal Studies, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

*Corresponding author:
fazilahrazali@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

The lack of studies evaluating CEFR-aligned English language curricula implemented in Malaysia is a serious concern impacting students' language learning environment. This issue becomes increasingly worrying in higher education because of the lack of curriculum standardization across various higher educational institutions. As students are the direct recipients of the implemented English curriculum, this study aimed to explore factors that influence university students' curriculum acceptance by examining their attitude, speaking anxiety, and remote learning readiness. It utilized a quantitative correlational design by distributing a validated and reliable cross-sectional survey to 405 undergraduate students at a Malaysian public university that recently underwent CEFR-aligned curriculum restructuring. Pearson's correlation tests and multiple linear regression analyses were conducted to examine the relationships between variables and determine their predicting effects. Findings indicated that students' attitude ($r=0.73$) and remote learning readiness ($r=0.67$) are significant predictors of curriculum acceptance. The regression model is statistically significant and has a strong model fit ($R^2=0.60$). These findings provide crucial students' perspectives for curriculum revisions and better implementation of the CEFR-framework in Malaysia. Insights from the study are applicable to CEFR-alignment policies, curriculum design, teaching, and assessment practices done in other ESL/EFL (English as a second/foreign language) countries around the world.

Keywords: CEFR evaluation, CIPP, curriculum evaluation, English language curriculum, student acceptance

1. Introduction

The English Language Education Reform: The Roadmap 2015-2025 is an integral English language education policy in Malaysia, and its principal component is the integration of Common European Framework of Reference for Languages (CEFR) into English language curricula across various education levels. This is due to the need for higher English language proficiency among Malaysian students especially in the highly competitive global market (Savski, 2021). The CEFR framework provides comprehensive guidelines in informing language learning, teaching, and assessment practices by providing detailed insights on communicative language activities, strategies, and competences.

The alignment of teaching and assessment practices to these CEFR guidelines has become a key component in English language teaching (ELT) practices in many countries worldwide. In Malaysia, the CEFR-aligned Standards-Based English Language Curriculum (SBELC) was introduced at primary and secondary schools across the country (Musa et al., 2021). This was followed by the development of CEFR-aligned textbooks, teaching materials, and national examinations (Zaki, 2021; Alias et al., 2021; Baharum et al., 2021). On a more global scale, the CEFR alignment efforts were seen in various forms like the integration into the national curriculum in Thailand (Chartrakul & Damnet, 2021), the development of students' 21st-century skills through curriculum design and assessment in Indonesia (Miqawati et al., 2023), and its application in shifting towards communicative language teaching in Finland and Poland (Aribas, 2021). These examples demonstrate the large-scale adoption of the CEFR framework in numerous ELT practices around the world.

Despite the widespread implementation of the CEFR framework in English language education, its effectiveness is rarely examined through systematic evaluation, particularly at the higher education level (Sahib & Stapa, 2021). While CEFR alignment aims to improve the quality of language education, many initiatives were implemented without robust mechanisms to assess their impacts on teaching, learning, and assessment practices. As Nurhayati (2024) notes, continuous monitoring and evaluation are essential for identifying strengths and weaknesses and for improving educational reforms. However, existing CEFR initiatives often lack structured evaluative frameworks that meaningfully involve key stakeholders, despite evidence that stakeholder engagement is critical for informed policy formulation and evaluation (Sharma & Kumar, 2023). While there are existing previous CEFR-related studies from school administrators' and teachers' perspectives, students' experiences and reception towards this curricular reform remain underexplored (Jamain, 2023).

This absence represents a significant evaluative gap in the current English language education reform in higher education. Addressing this gap through systematic curriculum evaluation is essential to ascertain the language education policy effectiveness and to support the success of CEFR-aligned reforms in higher education. Moreover, post COVID-19 pedagogical issues which significantly impacted students' cognition, behavior, and their learning environment make this curriculum evaluative gap even more critical. Concerns over increased anxiety level in CEFR-aligned speaking assessments, consistently high speaking anxiety during online learning, and the pressure of adapting to online English language classes among students may affect their attitude towards English language learning and their engagement with the CEFR-aligned curriculum implemented (Zulkiflee et al., 2023; Zulkiflee & Nimehchisalem, 2022; Mahmud & Mahmud, 2023; Ahmed et al., 2021). Therefore, to address the stated issues, the study aims:

1. To determine the relationship between university learners' attitude and their acceptance of CEFR-aligned English Language Curriculum
2. To determine the relationship between university learners' speaking anxiety and their acceptance of CEFR-aligned English Language Curriculum
3. To determine the relationship between university learners' remote learning readiness and their acceptance of CEFR-aligned English Language Curriculum
4. To determine the dominant predictor—among attitude, speaking anxiety, and remote learning readiness—of university learners' acceptance of a CEFR-aligned English language curriculum

Expanding on earlier CEFR studies in university contexts (Ng & Ahmad, 2021; Nawawi et al., 2021; Uri & Aziz, 2024), this study bridges the gap between school-based and higher education CEFR implementations. While previous Malaysian evaluations centered on secondary levels and teacher readiness (Uri & Aziz, 2018; Uri & Aziz, 2020; Ibrahim et al., 2023), this research extends the scope by sampling undergraduates and evaluating curriculum acceptance through online delivery. The focus on digital modality aligns with current technology-integration literature regarding curriculum adoption and learner experience (Mohammed et al., 2021; Fahrughazi et al., 2022; CH'NG et al., 2024). Hence, incorporating remote learning readiness serves as a meaningful extension of existing studies in the field. Crucially, this study addresses a distinct gap in the literature as no prior study has simultaneously examined attitude, speaking anxiety, and remote learning readiness as joint predictors of CEFR curriculum acceptance. By synthesizing these specific variables within the Malaysian higher education context, this study provides an original empirical framework that transcends the singular focus of previous research. It directly addresses a critical gap in Malaysian higher education research that has formerly focused on structural rather than affective or modality-specific factors (Uri & Aziz, 2018; Foley, 2019; CH'NG et al., 2024).

Learner attitude is included due to its strong correlation with oral performance and engagement in CEFR-aligned contexts (Wong & Ismail, 2023; Stognieva & Popov, 2021; Chou, 2018), while speaking anxiety acts as a proximal risk factor that specifically hinders communicative task uptake in online delivery modes (Mahmud & Mahmud, 2023; Teoh & Liao, 2021). Remote-learning readiness explains the practical capacity such as digital literacy and self-regulation which are required to navigate the shift toward online CEFR instruction, a determinant of perceived usefulness and success in contemporary programs (Almazroui & Alkaabi, 2022; Azarfam, 2022; Hermanto, 2015; Muttaqin & Ida, 2015). By synthesizing these three variables, the study examines whether acceptance is driven by internal dispositions or external technological readiness, aligning with recent literature that underscores the intersection of affect and technology in predicting students' curriculum acceptance (Marzaini & Yusoff, 2022). Theoretical implications of the study include contextual application of both the Social-Cognitive learning theory (Bandura, 1974) and curriculum evaluation model (Stufflebeam, 1971) by highlighting the crucial role of students as curriculum recipients in influencing CEFR-aligned curriculum acceptance. The study has crucial institutional and pedagogical implications as it enables curriculum leaders to make evidence-based curriculum development and design decisions (Legak and Wahi, 2020; Ramulumo, 2023). Overall, this study contributes to more effective student-centred and context-specific CEFR framework adoption in the global English language education.

2. Literature Review

2.1. Attitude in Acceptance of English Language Curriculum

Attitude refers to “a feeling or opinion one has about something and most often is reflected in one's behavior towards a construct. It often involves feelings and belief systems and the disposition to behave in each way” (Jamshed et al., 2023). Attitude is a key factor driving students' motivation and positive attitudes are often associated with increased classroom participation (Mumba, 2024). This is supported by Ismaili (2021) which claimed that attitude towards learning is a significant influencing factor in students' classroom engagement and overall success of English language acquisition. Aldujayn and Alsubhi (2020) stated that students' attitudes toward online language classes were found to be significantly affected by the usability and accessibility of the online learning platforms. Moreover, social factors such as the absence of peer interaction negatively affected students' attitude, leading to the lack of students' focus during online classes (Mukhtar et al., 2020). These studies demonstrated the sensitive nature of students' attitude which can either influence or be affected by various internal and external factors. Therefore, based on Bandura's (1974) Social Cognitive Theory, university learners' cognitive learning factor was operationalized as learners' attitude and the first hypothesis was developed.

H₁: There is a significant relationship between university students' attitude and their acceptance of CEFR-aligned English Language Curriculum.

2.2. Speaking Anxiety in Acceptance of English Language Curriculum

Speaking anxiety is “a major barrier to students' willingness to communicate” (Alaqeel & Altalhab, 2024), and it is a common issue that affects English language learners. The prevalence of speaking anxiety among students is still very high across many different contexts. This issue is not limited to only learners with low English language proficiency as even students with high language proficiency experience speaking anxiety (Hasibuan & Irzawati, 2020). The debilitating effects that speaking anxiety have towards learners and their learning experiences are well-documented in the literature. Dong (2023) found that there is a negative correlation between speaking anxiety and learners' self-efficacy. This implies that the probability of learners who demonstrate high levels of anxiety to engage in speaking activities is significantly lower, depriving them of speaking opportunities and hindering their English language acquisition (Dong, 2023). This finding is supported by Grieve et al. (2021) who found that learners' fears towards oral presentations and public speaking often disrupted their language learning experiences which resulted in speaking avoidance behaviors such as refraining from practicing their speaking skills. These mental and behavioral setbacks trap learners into a perpetual vicious cycle where anxiety hinders language practice, which then amplifies the speaking anxiety even further.

However, the context in which communication occurs is a key factor in managing this cycle. Research comparing speaking anxiety in online contexts versus face-to-face context has frequently indicated that comparably lower levels of apprehension were elicited by online environments (Rahmi & Murtafi'ah, 2022; Emara, 2025; Cheng & Sun, 2025). Studies involving English as a Foreign Language (EFL) echo this as synchronous voice-based and online speaking tasks repeatedly resulted in reduced anxiety and enhanced performance in comparison to traditional in-person interactions (Rahmi & Murtafi'ah, 2022; Bárkányi & Brash, 2025; Chew & Chen, 2021). However, these findings are nuanced and can be conflicting due to influential factors such as social presence and modality. While a virtual space tends to alleviate anxiety, a virtual audience can raise stress levels similar to levels in a physical classroom (Moïse-Richard et al., 2021). Drawing on Bandura's (1974) Social Cognitive Theory, which underscores the triadic reciprocal interaction between personal, environmental, and behavioural factors, university learners' behavioral learning factor was operationalized as learners' speaking anxiety and the second hypothesis was developed.

H₂: There is a significant relationship between university students' speaking anxiety and their acceptance of CEFR-aligned English Language Curriculum.

2.3. Remote Learning Readiness in Acceptance of English Language Curriculum

Remote learning readiness is the students' ability to “adapt their learning approaches rapidly to suit the needs of a virtual learning environment” (Tan et al., 2023). Following the COVID-19 pandemic, the global educational landscape drastically shifted in favor of remote learning over traditional physical classrooms. One of the most notable issues with remote learning is decreased motivation level and lower classroom participation among students. These could be attributed to technological issues such as internet connection and availability of electronic devices and the absence of social interactions between peers (Slykerman et al., 2022). Another study reported similar findings where it was argued that remote learning lacks the social proximity that is characteristic of traditional learning environments where students have face-to-face interactions in the classroom (Ceballos et al., 2021). Additionally, students can get direct support from teachers in the typical physical learning setting, but in remote learning settings, such physical support is essentially non-existent which leads to compromised learning experiences (Abuhammad, 2020). These potential adverse effects of remote learning towards students' language experiences must be addressed so that effective English language learning can still occur even

in an online environment. Therefore, based on Bandura's (1974) Social Cognitive Theory, university learners' environmental learning factor was operationalized as learners' remote learning readiness and the third hypothesis was developed.

H₃: There is a significant relationship between university students' remote learning readiness and their acceptance of CEFR-aligned English Language Curriculum.

2.4. Curriculum Acceptance of English Language Curriculum

Curriculum acceptance refers to "the views of all students who are still going through and have gone through this program" (Kalkan & Genç, 2023). The theoretical underpinning of students' curriculum acceptance is the CIPP (context, input, process, product) model of curriculum evaluation proposed by Daniel L. Stufflebeam in 1971. The CIPP model was adopted in this study due to its systematic approach to curriculum evaluation where language educators can gain insights into the students' reception towards the implemented CEFR-aligned English language curriculum at the university. Many previous studies also utilized the model in different contexts such as the evaluation of football school curriculum (Nugraha, 2022) and early childhood education during the COVID-19 pandemic (Suryana, 2023). In English Language Teaching (ELT), the model was also utilized extensively in evaluating effectiveness of language curricula such as identifying specific weaknesses (Manan et al., 2020; Yaşıtbaş & Kavgacı, 2020) and evaluating instructional alignment and student satisfaction (Tunç, 2010; Ulum, 2016). Collectively, these literatures affirm the CIPP model's suitability and universality to be used in different contexts, including the evaluation of CEFR-aligned English language curriculum carried out in this study.

The first component of the CIPP model is the context evaluation which explores the perceived curriculum relevance and allows educators to measure how well learners' expectations and personal learning goals are aligned with the implemented curriculum (Asiandrayani et al., 2021). Learners' attitude influences their acceptance of the curriculum context which is reflected in their perceptions towards the curriculum relevance, value, and appropriateness. Secondly, the input evaluation explores the perceived effectiveness of instructional materials and overall teaching approach employed by examining its adequacy, availability, and usefulness for learning purposes (Kang & Xu, 2024). Additionally, learners' remote learning readiness shapes their acceptance of the curriculum input because readiness allows learners to actively utilize audiovisual aids and online learning resources provided.

Next, the process evaluation explores the actual implementation of the curriculum, i.e. the teaching and learning process which allows educators to capture learners' experiences and reactions to the curriculum (Omid et al., 2020). Learners' speaking anxiety affects their acceptance of the curriculum process which is reflected in their participation in classroom activities and interaction with teachers and peers. Finally, the product evaluation explores the perceived educational effectiveness and attainment of learning outcomes as well as students' satisfaction level towards it (Indriyana et al., 2024). Collectively, these learner factors indirectly influence their acceptance of the curriculum product which is the overall perceived effectiveness of the CEFR-aligned curriculum implemented at the university.

Prior CEFR-related studies in Malaysia have highlighted notable tensions and gaps pertaining to stakeholder perceptions and curriculum implementations. A study that involved teachers as a key educational stakeholder established generally negative perceptions from teachers towards the CEFR framework (Nawawi et al., 2021). Policy-reality inconsistencies were discovered as teachers reported insufficient CEFR training and understanding which affected the process of CEFR implementation (Abidin & Hashim, 2021). These tensions were further amplified as Shak et al. (2021) questioned the suitability of imported CEFR-aligned English textbooks that are not culturally relevant for Malaysian students. Ironically, although many empirical studies had highlighted the urgent need for teachers'

professional development courses on the CEFR framework, intervention strategies by local educational leaders remained limited (Jamain & Jamaludin, 2023).

For instance, Samad and Haron (2021) found that despite some training on CEFR-aligned formative assessment methods, actual applications of those approaches in classroom practices were still low. This also demonstrates another significant gap which is difficulty aligning CEFR-aligned curriculum with classroom practices, evident in multiple prior studies. Zaki and Darmi (2021) claimed that many aspects of the curriculum developed using the CEFR framework were inadequately articulated or implemented in practice such as CEFR-based proficiency level measurement inconsistencies in universities (Baharum et al., 2021) as well as teacher-administered CEFR-Readiness Test design limitations (Marzaini & Yusoff, 2022). Trickleing down to curriculum recipients, these gaps were further reflected in students' learning experiences where CEFR-aligned speaking and writing tests presented setbacks such as barriers to higher language proficiency and psychological anxiety impacting effective learning (Ch'ng et al., 2024; Che Nor & Md Yunus, 2022). In summary, these studies highlight the crucial mismatch between policy aims for international standardizations and complex realities of the local Malaysian educational landscape. Therefore, based on Stufflebeam's (1971) CIPP model of curriculum evaluation, university learners' curriculum acceptance was operationalized, and the fourth hypothesis was developed.

H₄: There is a significantly dominant predictor—among attitude, speaking anxiety, and remote learning readiness—of university students' acceptance of CEFR-aligned English Language Curriculum.

2.5. Framework of the Study

The theoretical underpinning of this study is Bandura's (1974) Social Cognitive Theory where university learners' cognitive, behavioral, and environmental factors were examined as independent variables of the study which are attitude, speaking anxiety, and remote learning, respectively. These three independent variables were hypothesized to directly affect the dependent variable which is the students' acceptance of the CEFR-aligned English language curriculum. Further detailing the main constructs; Attitude was measured through cognitive attitude and emotional attitude sub-constructs; Speaking anxiety was measured through the sub-construct of interpersonal anxiety; Remote learning readiness was measured through self-directed learning and computer & internet efficacy sub-constructs. On the other hand, curriculum acceptance was measured through four sub-constructs of CIPP model of curriculum evaluation by Stufflebeam (1971) which are context, input, process, and product evaluations. Examining all these variables allow factors that affect university learners' learning predispositions to be measured and correlated against the level of their CEFR-aligned English language curriculum acceptance. The conceptual framework acts as guiding anchor of the study and allows the relationships between these variables to be analyzed and established clearly throughout the study period. The conceptual framework of this study is illustrated in Figure 1.

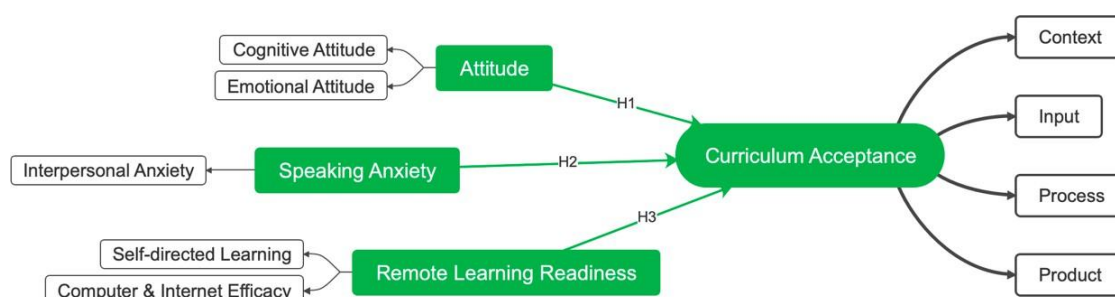


Figure 1. Conceptual Framework of Study

3. Methodology

The study employed a correlational research design which examined the relationships between attitude, speaking anxiety, remote learning readiness and curriculum acceptance. This cross-sectional survey was conducted at a public university in the Bangi district of Selangor, Malaysia. The selection of the university as the study location is because of its recent English language curriculum CEFR-aligned restructuring by adopting the ADDIE model (Musa et al., 2021). However, design limitations were therefore indicated; the cross-sectional design suggests that causal relationships between the variables must not be inferred and the study focus on a single higher education institution suggests that results should be interpreted with caution when generalizing to other populations or settings. Additionally, this study evaluated student acceptance instead of academic performance as it focuses on how the implemented CEFR-aligned curriculum influences students' learning experiences and their perceived curriculum effectiveness rather than evaluating performance measures. This study has obtained both the ethical approval from the Ethics Committee for Research Involving Human (JKEUPM) with reference number UPM.TNCPI.800-2/1/7 and voluntary informed consents from the participants.

3.1. Sampling

The study population is comprised of undergraduate students at the university. The study employed a two-step sampling method which entails purposive sampling followed by random sampling. First, the sampling frame was defined based on the inclusion criteria which is their enrolment in offered English courses, and it was obtained from the university's administration office. Based on the name lists of registered students, there is an estimated 5,705 total number of undergraduate students enrolled in English courses in semester 1. Based on this population size, estimated sample size in this study was determined by using Cohen's (1988) guidelines which stated that to detect a small effect size of 0.20 with 0.8 statistical power in a two-tailed correlation test at 0.05 significance level, an estimated minimum of 393 number of participants are required to be recruited in the study. This first sampling stage ensured that the participants have experienced the CEFR-aligned English language curriculum for at least one semester of English language learning at the university, and they can therefore provide relevant insights and feedback on the curriculum.

Following that, subject recruitment was done by randomly distributing the survey questionnaire to respondents at the end of semester 1 2024-2025 in January 2025 via their respective English language instructors. This approach leverages established instructor-student relationships (Nawai & Said, 2020) and allows the reaching of larger audience and higher response rates. A total of 600 students from all faculties across 20 class sections were invited to participate and 413 responses were completed, resulting in a response rate of 68.8% that met the predetermined minimum estimated sample size for correlation analyses. From 413 responses, 1 response was excluded due to the lack of informed consent, and 7 responses were removed as they were identified as outliers during the preliminary data cleaning stage, making the final number of responses analyzed to be 405 (n=405). By employing this two-step sampling method, only participants who have directly experienced the CEFR-aligned curriculum at the university were considered for selection and then selected, ensuring both contextually and statistically accurate procedure for data collection.

3.2. Instrumentation

The instrument employed in this study is a survey questionnaire that was adopted and adapted from previous studies. It comprises 68 items, each rated on a 5-point Likert scale ranging from 'Strongly Disagree' to 'Strongly Agree'. A 5-point scale was selected to minimize respondent fatigue and ensure clarity, as fewer response options can reduce the cognitive burden associated with complex decision-making (Aybek & Toraman, 2022). Demographic information, including respondents' faculty, MUET level, number of English courses completed, and year of study was collected and a mapping table

illustrating and linking each main construct to its source instrument and subscales is provided in Table 1.

Table 1. Construct-Instrument Mapping

Construct	Source	Sub-construct
Attitude →	Jamshed et al. (2023) →	Cognitive & Emotional (14 items)
Remote Learning Readiness →	Tan et al. (2023) →	Self-Directed Learning & Computer/Internet Efficacy (12 items)
Speaking Anxiety →	Alaqeel & Altalhab (2024) →	Interpersonal Anxiety (12 items)
Curriculum Acceptance →	Kalkan & Genç (2023) →	Context, Input, Process, Product (30 items)

3.3. Data Collection and Analysis

The content validity of the instrument was established by a panel of two subject matter experts, selected for their specialized expertise in the CEFR framework and the Malaysian higher education context. While larger panels were often discussed in psychometric literature, this study followed the evidence-based guidelines provided by Yusoff (2019), which established that a minimum of two experts is acceptable for content validation provided the Content Validity Index (CVI) meets a higher threshold of at least 0.80. The appointed experts reviewed the items on a 4-point scale of relevance, from which the CVI was calculated. In this study, both the S-CVI/Ave (scale-level content validity index based on the average method) and S-CVI/UA (scale-level content validity index based on the universal agreement method) validity indices were 1.0. This result indicates excellent validity and universal agreement across all survey items, exceeding the recommended thresholds for a two-expert panel.

Table 2. Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity

Construct	KMO	Bartlett	Component	Initial Eigenvalues	
				Total	% of Variance
Attitude	.92	<.001	1	6.09	43.94
			2	1.21	8.65
Remote Learning Readiness	.89	<.001	1	5.77	48.10
			2	1.36	11.29
Speaking Anxiety	.95	<.001	1	7.62	63.47
Curriculum Acceptance	.97	<.001	1	17.69	58.95
			2	1.56	5.19
			3	.94	3.13
			4	.88	2.92

Table 3. Factor Loadings

Attitude				
Item	Component 1	Component 2	Component 3	Component 4
A1		.78		
A2		.71		
A3		.63		
A4		.45		
A5		.49		
A6		.61		
A7		.45		

A8	.53			
A9	.65			
A10	.55			
A11	.93			
A12	.87			
A13	.62			
A14	.41			
Remote Learning Readiness				
RLR1		.51		
RLR2		.42		
RLR3		.81		
RLR4		.56		
RLR5		.79		
RLR6		.55		
RLR7	.52			
RLR8	.90			
RLR9	.90			
RLR10	.87			
RLR11	.76			
RLR12	.72			
Speaking Anxiety				
SA1	.76			
SA2	.79			
SA3	.80			
SA4	.88			
SA5	.87			
SA6	.81			
SA7	.84			
SA8	.76			
SA9	.84			
SA10	.68			
SA11	.77			
SA12	.75			
Curriculum Acceptance				
CA1				.74
CA2				.77
CA3				.48
CA4				.63
CA5				.79
CA6		.75		
CA7		.77		
CA8		.61		
CA9		.78		
CA10		.79		
CA11		.5		
CA12			.42	
CA13			.49	
CA14			.64	
CA15			.66	
CA16			.46	
CA17			.57	
CA18	.47			
CA19	.51			
CA20	.49			
CA21	.51			

CA22	.61			
CA23	.64			
CA24	.86			
CA25	.86			
CA26	.83			
CA27	.85			
CA28	.82			
CA29	.80			
CA30	.72			

To establish the construct validity of the instrument, exploratory factor analysis (EFA) using the Principal Component Axis (PCA) extraction method with Direct Oblimin rotation on each multi-item construct was conducted. Prior to that, suitability of the factor analysis was confirmed through the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity. The results in Table 2 confirmed the suitability of factor analysis as the Kaiser-Meyer-Olkin measure verified the sampling adequacy which ranged from .89 to .97 and there was a significant difference in the Bartlett's test ($<.01$), both meeting the recommended statistical thresholds (Hair et al., 2019). Subsequently, results from the factor analysis in Table 3 showed that all construct items loaded significantly ($>.40$) on their respective factors which confirmed the subscale structures (Hair et al., 2019).

The instrument's reliability was established by conducting a pilot study involving 30 students from the population which yielded Cronbach's alpha values of .7 for both attitude and remote learning readiness constructs and .9 for both speaking anxiety and curriculum acceptance constructs, both indicating acceptable and excellent reliability of items, respectively (Mohd Arof et al., 2018). IBM Statistical Package for the Social Sciences (SPSS) software was utilized to analyze data in this study due to its ability to handle complex statistical tests such as descriptive statistics and correlation analyses. Prior to conducting the statistical tests, the possibility of common method bias introduced in a single self-reported survey was addressed by conducting a Harman's single-factor test. The results confirmed that common method biases such as social desirability and acquiescence did not pose a significant threat to the validity of this study's findings because the first factor accounted for only 37.52% of the variance, which is well below the accepted threshold of 50% (Podsakoff et al., 2003).

Following that, data was further analyzed by conducting descriptive analysis to provide a general pattern of the collected data by reporting the mean scores and standard deviations along with its interpretations based on guidelines by Zaki & Ahmad (2017). Then, correlation analyses via Pearson's test and multiple linear regression were conducted to determine the relationships and the predicting effects of the independent variables towards the dependent variable. The strength of the Pearson's correlation coefficients was then interpreted based on guidelines by Cohen (1988). While composite Likert scores utilized in this study contain inherent measurement error, multiple linear regression was chosen over more advanced modeling approaches like structural equation modeling (SEM) due to its practical interpretability as its clear and interpretable coefficients are more applicable for broader educational policy and curricular decision-making.

4. Findings

4.1. Descriptive Analysis of Variables

Table 4. Summary of the Respondents' Demographic Information

		Frequency(f)	Percentage (%)
Faculty	FEP	13	3.2
	FFAR	5	1.2
	FKAB	52	12.8
	FGG	20	4.9

	FPEND	7	1.7
	FPER	23	5.7
	FPI	100	24.7
	FSK	34	8.4
	FSSK	1	.2
	FST	146	36.0
	FUU	3	.7
	CITRA	1	.2
MUET level	2 to 2.5	5	1.2
	3 to 3.5	197	48.6
	4 to 4.5	199	49.1
	5 to 5+	4	1.0
Year of Study	First year	165	40.7
	Second year	179	44.2
	Third year	42	10.4
	Fourth year	19	4.7
Number of English courses completed	One	225	55.6
	Two	131	32.3
	Three	39	9.6
	Four	10	2.5

Table 4 shows the summary of the respondents' demographic information. There were high number of respondents from Faculty of Science and Technology (FST, $f=146$, 36%) and Faculty of Islamic Studies (FPI, $f=100$, 24.7%), and Faculty of Engineering and Built Environment (FKAB, $f=52$, 12.8%). Moderate representations of respondents were observed from Faculty of Health Sciences (FSK, $f=34$, 8.4%), Faculty of Medicine (FPER, $f=23$, 5.7%), and Faculty of Dentistry (FGG, $f=20$, 4.9%). Low representations of respondents were recorded from Faculty of Economy and Management (FEP, $f=13$, 3.2%), Faculty of Education (FPEND, $f=7$, 1.7%), Faculty of Pharmacy (FFAR, $f=5$, 1.2%), Faculty of Law (FUU, $f=3$, .7%), Faculty of Social Science and Humanities (FSSK, $f=1$, .2%), and School of Liberal Studies (CITRA, $f=1$, .2%). The distribution analysis shows a faculty distribution skew and this imbalance may have introduced introduced a degree of systematic bias into the findings where the conclusions drawn may be more reflective of trends specific to these faculties rather than being fully representative of the broader population across all faculties.

In terms of the respondents' English language proficiency and their MUET level, most of them were Low Independent (MUET level 3 to 3.5, $f=197$, 48.6%) and High Independent (MUET level 4 to 4.5, $f=199$, 49.1%) students. There were very limited number of respondents with Basic English language proficiency (MUET level 2 to 2.5, $f=5$, 1.2%) as generally the university requires a minimum MUET level of 3 for university enrolment. Similarly, Proficient respondents (MUET level 5 to 5+, $f=4$, 1.0%) were also scarce in this study as this level of English language proficiency is only typically observed among English language teachers because English is not the country's first language.

Most of the respondents were first year ($f=165$, 40.7%) and second year students ($f=179$, 44.2%). There were limited numbers of third year ($f=42$, 10.4%) and fourth year students ($f=19$, 4.7%). Additionally, most of them were still in the early stage of their English language learning at the university where they had only completed one English course ($f=225$, 55.6%). Some of them had moderate level of experience with the curriculum with two English courses being completed ($f=131$, 32.3%) while smaller groups of other respondents had completed the whole English language curriculum at the university with three ($f=39$, 9.6%) and four ($f=10$, 2.5%) English courses being completed respectively.

4.2. Level of Students' Attitude

Table 5 summarizes the overall level of respondents' cognitive and emotional attitude towards English language learning.

Table 5. Level of Attitude

Construct	Sub-Construct	Mean	SD	Interpretation
Attitude	(Overall)	4.36	.42	Very High
	Cognitive Attitude	4.48	.43	Very High
	Emotional Attitude	4.24	.48	High

The findings established that the respondents strongly believe ($\bar{x}=4.48$, $s=.43$) in the importance of the English language for their education, confidence, and future career prospects. However, their own motivation level to learn English and their social language learning circle could be further improved ($\bar{x}=4.24$, $s=.48$) for more effective language learning experience. This phenomenon is similar with results from a study by Wati (2023) where the participants' high cognitive attitude contradicts their low affective attitude towards English language learning, implying that there is a crucial disparity between learners' cognitive beliefs and motivation level.

4.3. Level of Students' Speaking Anxiety

Table 6 summarizes the respondents' overall level of speaking anxiety during English classes at the university.

Table 6. Level of Speaking Anxiety

Construct	Sub-Construct	Mean	SD	Interpretation
Speaking Anxiety	(Overall)	2.97	.90	Moderate

The results highlighted that the respondents exhibited moderate signs of anxiety ($\bar{x}=2.97$, $s=.90$) such as trembling hands and shaking voice which stemmed from nervousness, fear, and panic when required to speak in English with either with the teacher or fellow classmates. However, it is important to note that these signs of anxiety did not prevent them from participating in lessons as more behavioral signs of anxiety such as avoiding group participations, hesitating when responding in English or deciding to not speak in English were found to be low. Majid and Mohamad (2024) explained that despite some ongoing anxiety about speaking, learners were able to remain active and confident during classroom lessons in the context of flipped classroom approach. Another similar study also affirmed that speaking anxiety does not necessarily lead to lower classroom participation because learners are able to develop coping mechanisms that enable them to overcome their fears and still speak in English (Wang, 2022).

4.4. Level of Students' Remote Learning Readiness

Table 7 below summarizes the respondents' overall level of remote learning readiness during online English classes.

Table 7: Level of Remote Learning Readiness

Construct	Sub-Construct	Mean	SD	Interpretation
Remote Learning Readiness	(Overall)	4.12	.47	High
	Self-directed Learning	3.94	.56	High
	Computer & Internet Efficacy	4.31	.48	Very High

Based on the findings, the respondents reported very high confidence ($\bar{x}=4.31$, $s=.48$) in using technological tools such as Microsoft Teams for online classes and web browsers to find information through any electronic devices like computer, tablets or smart phones. However, they had comparatively

lower self-directed learning abilities ($\bar{x}=3.94$, $s=.56$) and may require face-to-face English language classes and direct supervision from English language instructors due to virtual distractions such as social media and gaming during online learning settings. This research finding is consistent with results from multiple studies which established various reasons that could also deter effective online learning such as family distractions and unreliable internet connection (Maheshwari et al., 2022), trouble focusing and maintaining attention (Tan et al., 2022), and engagement with smart phones and social media during online lessons (Wang, 2022).

4.5. Level of Curriculum Acceptance

Table 8 below summarizes the overall level of respondents' CEFR-aligned English language curriculum acceptance at the university.

Table 8. Level of Curriculum Acceptance

Construct	Sub-Construct	Mean	SD	Interpretation
Curriculum Acceptance	(Overall)	4.29	.46	High
	Context	4.31	.50	Very High
	Input	4.29	.52	High
	Process	4.24	.50	High
	Product	4.32	.48	Very High

The findings showed positive learners' perceptions towards the overall curriculum relevance ($\bar{x}=4.31$, $s=.50$) and effectiveness ($\bar{x}=4.32$, $s=.48$). However, there is a need to improve the curriculum resources ($\bar{x}=4.29$, $s=.50$) and teaching-learning activities ($\bar{x}=4.24$, $s=.50$). This indicates a curriculum-pedagogical gap where the curriculum structure is widely perceived to be effective, but the implemented pedagogy which is the daily teaching-learning activities, the instructional strategies, and the supporting resources are perceived to be lagging behind. Without addressing these instructional gaps, the positive perceptions of the overall curriculum may eventually be undermined by a frustrating classroom experience. However, this tension can be resolved by incorporating audiovisual materials that can attract learners' interest and actively engage all their language skills. This is because interactive linguistic multimedia resources such as videos, pictures and animated content can provide clear instructions and engaging content (Rahmanu & Santosa, 2022) while facilitating cognitive growth (Yi & Wei, 2022). Similarly, there was a demand for learning activities that are interesting while encouraging peer interaction and group work. Multiple previous studies have established the role of cooperative learning experiences in fostering a sense of language learning community and support (Peter & Asmawi, 2023), providing learners with opportunities to practice language skills in context (Sudewi & Fadilah, 2023), and creating enjoyable learning experiences which leads to increased motivation (Kanellopoulou et al., 2021).

In terms of classroom activities, the respondents indicated that more activities should be provided so that they have adequate practice of the language skills learnt. This is because learners who are supplied with lots of opportunities to use their target language can attain higher language proficiency (Tran et al., 2021). Similarly, the results also implied that the classroom activities at the university were not sufficient to guarantee improvements in their English language proficiency. One key strategy that can be employed is through repetition of difficult course content until learners achieve mastery of language skills. Kriukova et al. (2024) found that spaced repetition, which is reviewing difficult content at increasing intervals, provide opportunities for learners to practice language skills frequently and prevent cognitive overload, leading to improved retention of vocabulary and grammar skills. Resulting from repeated practice of language skills, learners gain more confidence and attain better language fluency over time, enabling them to use their language skills effectively in various real-life contexts (Lo, 2024).

Revisions of specific language content were also indicated from the survey responses as the participants believe that at the end of the English courses, their writing, grammar, and overall English language

proficiency might not have had profound improvements. This implies that despite the focus on speaking and communicative competence, English language curricula at higher education institutions should not completely disregard the significance of other language skills in their curriculum content. This is because the mastery of academic writing skills is essential for various tasks that university students are expected to undertake (Hass & Lenong, 2021). Similarly, improving learners' grammatical accuracy allows them to develop metalinguistic knowledge and skills to engage multidisciplinary texts more analytically and critically (Tuttle & Munalim, 2022), which is paramount in the context of higher education. In contrast, university students who lack these foundational skills may struggle to meet academic standards and transition to the professional world, where effective written communication is equally valued.

4.6. Factors Influencing Curriculum Acceptance

Table 9. Skewness and Kurtosis of Variables

Variables	Skewness	Kurtosis
Attitude	-.068	-.817
Speaking Anxiety	.079	-.621
Remote Learning Readiness	.331	-.578
Curriculum Acceptance	.377	-1.146

Table 9 shows that the skewness and kurtosis values for all variables fall within the recommended range for normality (Hair et al., 2022), indicating that the dataset is approximately normally distributed. Furthermore, scatterplots for each predictor against the dependent variable revealed fitted lines that closely approximated a straight line, satisfying the assumption of linearity. Adherence to these parametric assumptions justified the use of Pearson's correlation and multiple regression analyses.

Table 10. Inter-predictor Correlation Matrix

Variable	Attitude	Remote Learning Readiness	Speaking Anxiety	Curriculum Acceptance
Attitude	1	.63**	-.19**	.73**
Remote Learning Readiness	.63**	1	-.25**	.67**
Speaking Anxiety	-.19**	-.25**	1	-.20**
Curriculum Acceptance	.73**	.67**	-.20**	1

** . Correlation is significant at the 0.01 level (2-tailed).

Results from Pearson's correlation test in Table 10 show that there was a statistically significant strong and positive relationship between respondents' attitude and their curriculum acceptance ($r=.73$, $p<.001$), accounting for 52.9% of the bivariate variance explained. This indicates that learners' attitude is substantially correlated to their curriculum acceptance, providing a strong statistical ground for the acceptance of the first hypothesis (H_1). On the other hand, although speaking anxiety was found to be significantly associated with curriculum acceptance ($p<.001$), the relationship was observed to be weak and negative ($r=-.20$), explaining only 4.1% of the bivariate variance. This suggests that while lower speaking anxiety is slightly linked to higher curriculum acceptance, allowing the second hypothesis (H_2) to be supported, its explanatory power is still very limited. In addition, remote learning readiness demonstrated a statistically significant strong and positive relationship with curriculum acceptance ($r=.67$, $p<.001$), explaining 44.4% of the bivariate variance. This result provides strong justification for the third hypothesis (H_3) to be accepted. Additionally, it should be noted that these figures represent bivariate variance explained in isolation and because the predictors shared significant variance, these values cannot be interpreted as independent or additive contributions to the multi-predictor regression

model. Overall, the findings highlighted higher influence of cognitive and environmental learning factors towards curriculum acceptance compared to behavioural factors.

4.7. Predictors of Curriculum Acceptance

Prior to conducting the multiple linear regression analysis, the presence of multicollinearity among independent variables was assessed by examining the correlation coefficients, Tolerance values, and Variance Inflation Factor (VIF) values. In this study, Pearson's correlation values among predictors remained below .7, Tolerance values were above .1, and VIF values were less than 10; collectively, these results confirmed the absence of multicollinearity (Brun et al., 2019; Hair et al., 2019). Subsequently, the multiple linear regression was performed to identify the dominant predictors of university students' acceptance of the CEFR-aligned English Language Curriculum. Diagnostic checks of the residuals were conducted where visual inspection of the residuals scatterplot indicated homoscedasticity and linearity, while the Normal P-P plot confirmed the normality of residuals.

Table 11. Regression Coefficients

Variable	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	.530	.179		2.962	.003	.178	.883
Attitude	.554	.045	.508	12.444	<.001	.467	.642
Remote Learning Readiness	.334	.041	.341	8.236	<.001	.254	.413
Speaking Anxiety	-.010	.017	-.019	-.597	.551	-.43	.023

The results in Table 11 indicated significant predicting effects of attitude and remote learning readiness towards curriculum acceptance because $p < .001$. On the other hand, the study discovered the non-predictive role of speaking anxiety towards curriculum acceptance because $p = .551$. This finding is consistent with the previously established weak correlation between speaking anxiety and curriculum acceptance. In terms of predictive strength, university students' attitude emerged as the dominant predictor with $B = .554$, followed by remote learning readiness with $B = .334$. The predicting effect of speaking anxiety can be considered negligible, as indicated by its non-significant beta coefficient ($B = -.010$, $p = .551$). This suggests that speaking anxiety does not exert a meaningful direct influence on curriculum acceptance when considered alongside other variables. Furthermore, the weak correlation strength indicates that anxiety accounts for only a minimal portion of the variance, making its unique contribution statistically insignificant in the presence of stronger predictors such as attitude and remote learning readiness. While speaking anxiety demonstrated a limited direct influence in the current model, its role may be more complex than a simple linear relationship. A plausible, albeit speculative, explanation is that speaking anxiety influences curriculum acceptance through indirect pathways, potentially by shaping learners' attitude or their remote learning readiness and this can be confirmed in future studies that employ formal mediation analyses.

Table 12. ANOVA Analysis

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	50.863	3	16.954	200.474	<.001 ^b
Residual	33.913	401	.085		
Total	84.777	404			

Table 13. Regression Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
.775 ^a	.600	.597	.29081	2.013

Subsequently, the regression model was found to be statistically significant, which is reported from Table 12 as $F(3,401) = 200.474$, $p < .001$. The very large F-value of 200.474 implied that collectively, the independent variables can predict university learners' curriculum acceptance accurately and the probability of this predicting effect to occur by chance is only less than .001. This is supported by the high adjusted R^2 value of .597 which is the variation explained by the model in the population, shown in Table 13 above. The adjusted R^2 value is interpreted as more than half (59.7%) of the variability in university students' CEFR-aligned English language curriculum acceptance is explained by the predictors. Furthermore, the Durbin-Watson value of 2.013 strengthened the viability of this model, proving that there is no autocorrelation in the residuals (errors between predicted and actual values) and the absence of Type I errors. In summary, alternative hypothesis (H_4) was accepted. The equation for the regression model is:

$$\text{Curriculum Acceptance} = 0.530 + 0.554\text{Attitude} + 0.334\text{Remote Learning Readiness}$$

5. Discussion

The collective findings of this study suggest that the successful implementation of the CEFR-aligned English language curriculum from students' perspectives is fundamentally predicated on an 'attitude-remote learning readiness' nexus rather than being hindered by emotional barriers like speaking anxiety. While the current CEFR language education policy aims to standardize proficiency, the results indicate that student acceptance is less about the linguistic content and more about the learner's psychological predisposition and digital learning environment. A critical interrogation of the findings reveals that student attitude acts as a primary gatekeeper for the successful implementation of new English language education policies; university students who hold positive attitudes are significantly more likely to respond well to the newly implemented curriculum structure. This suggests that for the CEFR alignment to move beyond a top-down mandate and achieve bottom-up adoption, students must first perceive the curriculum as relevant to their personal learning goals (Tan et al., 2021). As noted in previous literature, such attitudes are essential for improving learning engagement (Ramzan et al., 2023) and increasing the motivation required for active participation (Doley, 2023). However, the evidence further demonstrates that attitude does not function in isolation. The strong inter-predictor correlation between attitude and remote learning readiness highlights that curriculum acceptance is heavily influenced by digital self-efficacy. For the CEFR framework to be effectively implemented in the current digital age, students must feel technologically competent in the medium of delivery, as digital resource utilization forms the backbone of the modern online learning experience (Han et al., 2021). In the absence of this remote learning readiness, even a positive attitude may be insufficient to overcome the virtual challenges and distractions that necessitate high levels of self-regulation and independent learning behaviours (Araka et al., 2020).

On the other hand, while speaking anxiety demonstrated a significant bivariate correlation with curriculum acceptance, its failure to emerge as a unique predictor suggests that speaking anxiety is a redundant variable when accounted for alongside attitude and readiness, within the context of this study. Essentially, anxiety seems to be a sign of low readiness or a negative attitude rather than an independent factor of curriculum rejection. This implies that policy interventions focusing solely on anxiety reduction may be misplaced; instead, CEFR-aligned language curricula should prioritize building online self-efficacy and self-regulated learning behaviors (Ulfatun et al., 2021). When students are technologically prepared and intrinsically motivated, they possess the coping skills and learning drive necessary to maintain high engagement levels, effectively rendering their speaking anxiety a secondary concern (Gurajena et al., 2021; Tan et al., 2023). As a result, they are able to engage with the curriculum actively, leading to higher academic performance (Razali et al., 2022; Almufareh, 2021). This suggests that the broader aims of the CEFR language education policy should not simply focus on curricular and

pedagogical changes but should also consider the digital and psychological learning ecosystem where the learner feels both capable and willing to adapt to new language learning environments (Tan et al., 2023). This is because much of the variance anxiety explains is captured by more dominant cognitive and environmental factors, implying that while speaking anxiety is a relevant factor, it does not provide an independent contribution to curriculum acceptance beyond attitude and remote learning readiness.

This finding represents a theoretically crucial contribution that disputes the traditional assumption of anxiety as a common hindrance to learner engagement in language classes (Tsiplakides & Keramida, 2009). This outcome aligns with an emerging body of evidence, including the work of Hasanah and Desvitasari (2023) and Nabilah (2021), which reported non-significant relationships between anxiety and performance or learner beliefs (Rajani Na Ayuthaya & Sitthitikul, 2016; Krishnasamy, 2015). In lieu of being a stable predictor, anxiety is increasingly regarded as a context-dependent factor within a complex dynamic system where pedagogical factors such as task structure, scaffolded feedback, and instructional design play a more proximal role in forming learner's response to a curriculum (Wang et al., 2025; Korucu-Kis & Sanal, 2020; Nguyen, 2024). These approaches enable students to effectively employ anxiety coping mechanisms such as using task-based learning approaches (Badrasawi et al., 2020), focusing on communicative competence over language accuracy (Cabansag, 2020), and seeking for institutional support (Aziz et al., 2023).

The curriculum acceptance model developed in this study provides practical implications for CEFR-alignment initiatives. First and foremost, it highlights the significance of measuring and evaluating learners' receptions to the implemented curriculum for future curriculum revisions (Doley, 2023). Evaluating student feedback acts as an effective tool for identifying curriculum weaknesses (Bhurtyal et al., 2022) as integrating students' perspectives into the curriculum improves academic and emotional competencies by increasing a sense of belonging (Nirmalasari et al., 2024). Since attitude is the strongest predictor of student acceptance in this study, curriculum designers must prioritize cultivating positive students' emotional attitudes by creating interesting and enjoyable curriculum content (Ramzan et al., 2023). Additionally, pedagogical approaches that promote self-regulated learning skills are essential, allowing students to become proactive language learners who can effectively manage distractions in the virtual learning environment (Chui, 2024).

To translate these findings into policy enactment and curricular actions, the Ministry of Higher Education (MoHE) and university curriculum committees should supplement existing CEFR alignment strategies by adopting Key Performance Indicators (KPIs) that regularly measure the Student Acceptance Index (SAI). This can be carried out through regular end-of-semester student evaluations where institutions should be aiming for a minimum of 80% registered students reporting high curriculum acceptance (Likert score 4.0). By explicitly operationalizing SAI as a formal KPI, curriculum developers can identify 'problematic' language courses where low acceptance serves as a lead indicator for possible future classroom engagement issues. Secondly, instead of viewing post-COVID 19 online learning and language curriculum as separate unrelated factors, curriculum developers should actively address the direct impacts of the virtual environment toward language learning and overall language acquisition. Diagnostic assessment of students' remote learning readiness should be conducted prior to their enrolment in language courses so that corresponding interventions such as digital-readiness workshops can be planned in advance. This KPI is a direct policy enactment of the findings where technical and self-regulation barriers are actively addressed. In summary, by regularly measuring and evaluating students' learning attitude and digital self-efficacy through these KPIs, policymakers and curriculum leaders can mitigate the underlying causes of student resistance and ensure that the CEFR-framework implementation achieves sustainable, bottom-up acceptance across all Malaysian HEIs.

6. Conclusion

The study extends conceptual understanding of student acceptance of implemented English language curriculum at a HEI by integrating cognitive and environmental factors within the CEFR framework. By establishing that attitude and remote learning readiness as key predictors of curriculum acceptance, the study highlights the utmost necessity for HEIs to design learner-centred and context-sensitive CEFR-aligned English language curricula. On the other hand, the non-predictive role of speaking anxiety in this study offers a meaningful contribution as it challenges scholarly assumptions about its centrality in English language curriculum and instruction, especially in similar educational contexts. As the CEFR framework continues to be the prime source of reference for English language education, the study's evaluative contributions enable other HEIs around the world to measure the effectiveness of their own implemented English language curricula. Additionally, the study provides crucial contributions towards curriculum developers in redesigning and improving the English language curriculum and ensuring that they are well-aligned within the CEFR framework.

There are several limitations of the study. First, it evaluates the English language curriculum at only one public Malaysian university which implies that the results may not reflect relationships present in other HEI settings. Secondly, speaking anxiety was only measured through its interpersonal anxiety sub-construct and therefore the full construct may yield different results. Moreover, relationships between the variables across different subgroups such as proficiency levels were not further explored due to data limitations. These limitations provide clear pathways for future studies that could further examine, explain and fill in the gaps identified. First and foremost, future studies should evaluate other CEFR-aligned English language curricula offered at other public and private higher education institutions in Malaysia. By doing so, a more holistic monitoring of the CEFR-alignment initiatives and evaluation CEFR-aligned curricula can be ensured. Secondly, future studies that critically examine the influence of subgroups (i.e. faculties, proficiency levels) as well as potential mediating effects of attitude and remote learning on speaking anxiety may offer deeper insights into these observed relationships. Finally, mixed-methods curriculum evaluation studies can further supplement the findings in this study by exploring the issues from multiple perspectives through data triangulation, providing a more nuanced understanding of student acceptance of CEFR-aligned curricula in Malaysian HEIs.

References

- Abidin, N. Z. and Hashim, H. (2021). Common European Framework of Reference (CEFR): A review on teachers' perception & plurilingualism. *Creative Education*, 12(04), 727-736. <https://doi.org/10.4236/ce.2021.124051>
- Abuhammad, S. (2020). Barriers to distance learning during the COVID-19 outbreak: A qualitative review from parents' perspective. *Heliyon*, 6(11), e05482. <https://doi.org/10.1016/j.heliyon.2020.e05482>
- Ahmed, S. A., Othman, B. J., Gardi, B., Sabir, B. Y., Ismael, N. B., Hamza, P. A., Sorguli, S., Aziz, H. M., Ali, B. J., & Anwar, G. (2021). Students' attitudes towards learning English in the Kurdistan region of Iraq. *International Journal of English Literature and Social Sciences*, 6(3), 72–87. <https://doi.org/10.22161/ijels.63.11>
- Alaqeel, S., & Altalhab, S. (2024). Levels and major causes of Saudi students' speaking anxiety (SA) in EFL classrooms. *Theory and Practice in Language Studies*, 14(3), 771–780. <https://doi.org/10.17507/tpls.1403.18>
- Aldujayn, N., & Alsubhi, A. (2020). Saudi EFL teachers' interpretation toward creativity. *English Language Teaching*, 13(6), 162–170. <https://doi.org/10.5539/elt.v13n6p162>
- Alias, M., Raof, A. H. A., & Abdullah, T. (2021). Teaching CEFR-aligned writing to young learners: Practices and voices of teachers. *Pertanika Journal of Social Sciences and Humanities*, 29(S3), 229–246. <https://doi.org/10.47836/pjssh.29.s3.18>
- Almazroui, K., & Alkaabi, M. (2022). Preparing students for studying in a non-native language in math and science. In *International conference on multidisciplinary research* (pp. 114–129). <https://doi.org/10.26803/myres.2022.09>
- Almufareh, M. F. (2021). The impact of gamification and individual differences on second language learning among first-year female university students in Saudi Arabia. *Simulation & Gaming*, 52(6), 715–734. <https://doi.org/10.1177/10468781211056511>
- Araka, E., Maina, E., Gitonga, R., & Oboko, R. (2020). Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—Systematic review (2008–2018). *Research and Practice in Technology Enhanced Learning*, 15(1), 6. <https://doi.org/10.1186/s41039-020-00129-5>
- Aribaş, Ü. (2021). The contribution of CLIL approach to teaching English as a foreign language: Example of Finland, Poland, and Turkey. *Journal of History School*, 50, 438–464. <https://doi.org/10.29228/joh.48386>
- Asiandrayani, A., Sukardi, S., & Wilian, S. (2021). Performance evaluation of public elementary school supervisors in Maharam city during COVID-19. *Jurnal Ilmiah Profesi Pendidikan*, 6(3), 374–379. <https://doi.org/10.29303/jipp.v6i3.228>
- Aybek, E. C., & Toraman, Ç. (2022). How many response categories are sufficient for Likert-type scales? An empirical study based on the item response theory. *International Journal of Assessment Tools in Education*, 9(2), 534–547. <https://doi.org/10.21449/ijate.1132931>
- Azarfam, A. A. Y. (2022). *A comparative study of the Iranian and Malaysian English learners' communication anxiety, communicative competence, willingness to communicate (WTC), and international posture*. Research Square. <https://doi.org/10.21203/rs.3.rs-1842980/v1>

- Aziz, A., Rashid, R., & Zainudin, W. (2018). The enactment of the Malaysian Common European Framework of Reference (CEFR): National master trainer's reflection. *Indonesian Journal of Applied Linguistics*, 8(2), 452–460. <https://doi.org/10.17509/ijal.v8i2.13307>
- Badrasawi, K. J. I., Solihu, A., & Tunku Ahmad, T. B. (2020). Second language speaking anxiety among Malaysian postgraduate students at a faculty of education. *International Journal of Education and Literacy Studies*, 8(2), 54–60. <https://doi.org/10.7575/aiac.ijels.v.8n.2p.54>
- Baharum, N. N., Ismail, L., Nordin, N., & Razali, A. B. (2021). Aligning a university English language proficiency measurement tool with the CEFR: A case in Malaysia. *Pertanika Journal of Social Sciences and Humanities*, 29(S3), 125–140. <https://doi.org/10.47836/pjssh.29.s3.09>
- Bárkányi, Z., & Brash, B. (2025). Foreign language speaking anxiety online: Mitigating strategies and speaking practices. *ReCALL*, 37, 1–20. <https://doi.org/10.1017/S0958344025000060>
- Bhurtyal, K. M., Karki, S., Paudel, G., Baral, S. R., Gautam, A., Chaurasiya, R. K., & Khanal, U. (2022). Evaluation of students' perceptions and feedback about teaching-learning pharmacology in a medical college of central Nepal. *Journal of Chitwan Medical College*, 12(1), 111–115. <https://doi.org/10.54530/jcmc.668>
- Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., & Zimmermann, N. E. (2019). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, 47(1), 130–142. <https://doi.org/10.1111/jbi.13734>
- Cabansag, J. N. (2020). Speaking anxiety, English proficiency, affective and social language learning strategies of ESL engineering students in a state university in Northern Luzon, Philippines. *International Journal of English Linguistics*, 10(1), 372–386. <https://doi.org/10.5539/ijel.v10n1p372>
- Ceballos, M., Vitale, T., & Gordon, W. R. (2021). Remote continuity of learning and the COVID-19 pandemic: Educators' self-perceptions of preparedness. *Journal of Pedagogical Sociology and Psychology*, 3(2), 75–89. <https://doi.org/10.33902/jpsp.2021271304>
- Ch'ng, L., Mahmud, F. F., Sahari, S. H., Arshad, A., & Gafar, S. Z. (2024). Evaluating students' views on the importance and uselessness of CEFR in speaking test. *Issues in Language Studies*, 13(1), 166–180. <https://doi.org/10.33736/ils.6219.2024>
- Chartrakul, K., & Damnet, A. (2021). Role of the CEFR and English teaching in Thailand: A case study of Rajabhat universities. *Advances in Language and Literary Studies*, 12(2), 82–90. <https://doi.org/10.7575/aiac.alls.v.12n.2.p.82>
- Che Nor, W. N. and Md Yunus, M. (2022). ESL Writing anxiety in the CEFR setting: Does it affect the students?. *International Journal of Academic Research in Progressive Education and Development*, 11(3). <https://doi.org/10.6007/ijarped/v11-i3/14357>
- Cheng, Z., & Sun, P. P. (2025). Does L2 speaking anxiety differ in classroom and synchronous online learning environments? Evidence from EFL college students. *Journal of Computer Assisted Learning*, 41(5), e70115. <https://doi.org/10.1111/jcal.70115>
- Chew, S. Y., & Chen, Y. (2021). Speaking performance and anxiety levels of Chinese EFL learners in face-to-face and synchronous voice-based chat. *Journal of Language and Education*, 7, 43–57. <https://doi.org/10.17323/jle.2021.11878>

- Chou, M. (2018). Speaking anxiety and strategy use for learning English as a foreign language in full and partial English-medium instruction contexts. *TESOL Quarterly*, 52(3), 611–633. <https://doi.org/10.1002/tesq.455>
- Chui, Y. (2024). Research on Chinese university students' needs, goal-orientation, and self-regulated strategies in an ESL class. *Adult and Higher Education*, 6(3), 180–192. <https://doi.org/10.23977/aduhe.2024.060318>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Doley, S. K. (2023). Stimulus appraisal-based L2 attitude and motivation among Indian ESL learners. *International Journal of Instruction*, 16(2), 603–622. <https://doi.org/10.29333/iji.2023.16232a>
- Dong, W. (2023). Anxiety and self-efficacy in foreign language learning speaking among Chinese students. *Lecture Notes in Education Psychology and Public Media*, 27(1), 285–289. <https://doi.org/10.54254/2753-7048/27/20231231>
- Emara, E. (2022). A comparative and correlational study of face-to-face and face-to-screen EFL students' speaking anxiety with reference to complex dynamic systems theory. *Journal of Faculty of Education (Benha University)*, 33, 59–90. <https://doi.org/10.21608/jfeb.2022.301265>
- Fahrughazi, S., Omar, H., & Mohamad, M. (2022). Perceptions of postgraduates majoring in English education on speaking assessment module for online distance learning. *Theory and Practice in Language Studies*, 13(1), 78–88. <https://doi.org/10.17507/tpls.1301.10>
- Foley, J. (2019). Issues on the initial impact of CEFR in Thailand and the region. *Indonesian Journal of Applied Linguistics*, 9(2). <https://doi.org/10.17509/ijal.v9i2.20233>
- Grieve, R., Woodley, J., Hunt, S., & McKay, A. (2021). Student fears of oral presentations and public speaking in higher education: A qualitative survey. *Journal of Further and Higher Education*, 45(9), 1281–1293. <https://doi.org/10.1080/0309877X.2021.1948509>
- Gurajena, C., Mbunge, E., & Fashoto, S. G. (2021). Teaching and learning in the new normal: Opportunities and challenges of distance learning amid COVID-19 pandemic. *International Journal of Education and Teaching*, 1(2), 9–15. <https://doi.org/10.51483/ijedt.1.2.2021.9-15>
- Hair, J. F., Babin, B. J., Anderson, R. E., & Black, W. C. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.
- Han, S., Eum, K., Kang, H. S., & Karsten, K. (2021). Factors influencing academic self-efficacy among nursing students during COVID-19: A path analysis. *Journal of Transcultural Nursing*, 33(2), 239–245. <https://doi.org/10.1177/10436596211061683>
- Hasanah, A. P. U., & Desvitasari, D. (2023). Correlation between speaking anxiety and speaking achievement. *Journal of English Education*, 4(1). <https://doi.org/10.52333/djoe.v4i1.71>
- Hasibuan, A., & Irzawati, I. (2020). Students' speaking anxiety on their speaking performance: A study of EFL learners. In *Proceedings of the 3rd International Conference on Education, Science, and Technology* (pp. 67–70). <https://doi.org/10.2991/assehr.k.200115.017>

Hass, A. B., & Lenong, B. (2021). Assessing the academic writing skills of final year English second language (ESL) education students to determine their preparedness as language teachers: A practical approach at a university of technology. In *Education and new developments 2021* (pp. 371–375). InScience Press. <https://doi.org/10.36315/2021end079>

Hermanto, H. (2015). Typical errors in the writing of students participating in Mawapres selection in ITS Surabaya. *Nobel Journal of Literature and Language Teaching*, 6(1), 61–70. <https://doi.org/10.15642/nobel.2015.6.1.61-70>

Ibrahim, S., Aziz, A., Said, N., & Ismail, H. (2023). Malaysian English language teachers' willingness, readiness, needs and wants to develop graphic oral history ELT materials. *International Journal of Learning, Teaching and Educational Research*, 22(5), 101–118. <https://doi.org/10.26803/ijlter.22.5.5>

Ismaili, Y. (2021). Evaluation of students' attitude toward distance learning during the pandemic (Covid-19): A case study of ELTE University. *On the Horizon*, 29, 17-30. <https://doi.org/10.1108/OTH-09-2020-0032>

Indriyana, K. P., Saputri, R., Herawan, E., & Rusdiyana, R. (2024). Evaluation of differentiated learning in economics subjects using the CIPP model. *International Journal of Education, Social Studies, and Management*, 4(2), 244–256. <https://doi.org/10.52121/ijessm.v4i2.229>

Jamain, S., & Jamaludin, K. (2023). Issues and challenges in the implementation of the Common European Framework of Reference (CEFR) in Malaysian primary school. *International Journal of Academic Research in Business and Social Sciences*, 13(12), 721–735. <https://doi.org/10.6007/ijarbss/v13-i12/20394>

Jamshed, M., Alam, I., Hussain, N., & Banu, S. (2023). Investigating the attitudes of ESL learners toward the English language. *World Journal of English Language*, 13(8), 575–584. <https://doi.org/10.5430/wjel.v13n8p575>

Kalkan, L., & Genç, G. (2023). Evaluation of the intensive English language teaching programme for the fifth grade according to students' views. *International Electronic Journal of Elementary Education*, 15(5), 521–534. <https://doi.org/10.26822/iejee.2024.325>

Kanellopoulou, C., Pergantis, M., Konstantinou, N., Kanellopoulos, N. G., & Giannakouloupoulos, A. (2021). Foreign language web-based learning by means of audiovisual interactive activities. *Journal of Software Engineering and Applications*, 14(6), 207–232. <https://doi.org/10.4236/jsea.2021.146013>

Kang, L., & Xu, X. (2024). Progress and prospect of digital transformation of higher education in China: Bibliometric and visualized analysis. *Journal of Adult and Continuing Education*, 31(1), 109–134. <https://doi.org/10.1177/14779714241300086>

Korucu-Kis, S., & Sanal, F. (2020). Bridging in-class and out-of-class learning through podcast tasks. *International Online Journal of Education and Teaching*, 7(2), 636–653. <https://www.iojet.org/index.php/IOJET/article/view/802>

Krishnasamy, J. (2015). ESL students' anxiety in learning and communicating in the English language. *International Journal of English Language, Literature and Humanities*, 2(1), 268–282.

Kriukova, Y., Yamshynska, N., Купенок, Н., & Meleshko, I. (2024). Enhancing ELS students' language proficiency through spaced repetition. *Education. Innovation. Practice*, 12(8), 34–41. <https://doi.org/10.31110/2616-650x-vol12i8-005>

- Legak, T. C., & Wahi, W. (2020). Communicative language tasks to enhance young learners' communicative competence. *International Journal of Academic Research in Business and Social Sciences*, 10(6), 350–364. <https://doi.org/10.6007/ijarbss/v10-i6/7308>
- Lo, S. (2024). Vocabulary learning through viewing dual-subtitled videos: Immediate repetition versus spaced repetition as an enhancement strategy. *ReCALL*, 36(2), 152–167. <https://doi.org/10.1017/s0958344024000053>
- Mahmud, N. N., & Mahmud, N. (2023). Language learning anxiety in online learning: Malaysian ESL undergraduates' perspectives. *ELTALL: English Language Teaching, Applied Linguistics and Literature*, 4(2), 48–66. <https://doi.org/10.21154/eltall.v4i02.5766>
- Majid, N. A. A., & Mohamad, M. (2024). The effects of flipped classroom on Form 4 ESL learners' speaking performance in a Malaysian secondary school. *SHS Web of Conferences*, 182, 01008. <https://doi.org/10.1051/shsconf/202418201008>
- Manan, A., Fadhilah, M. A., & Kamarullah. (2020). Evaluating paper-based TOEFL preparation program using the CIPP model. *Studies in English Language and Education*, 7(2). <https://doi.org/10.24815/siele.v7i2.16467>
- Marzaini, A. F. B. M. M. M. and Yusoff, S. M. (2022). Assessing CEFR-Readiness test in Malaysian ESL classroom: An insight from English language teachers in Pulau Pinang. *Asian Journal of Assessment in Teaching and Learning*, 12(1), 1-9. <https://doi.org/10.37134/ajatel.vol12.1.1.2022>
- Miqawati, A., Wijayanti, F., & Ismailia, T. (2023). Adopting CEFR for English language learning in Indonesia to support 21st century skills. In *Proceedings of the 3rd International Conference on Social Science, Humanity and Public Health (ICoSHIP 2022)* (pp. 25–30). <https://doi.org/10.4108/eai.5-11-2022.2326530>
- Mohammed, N., Raof, A., & Yusof, M. (2021). Challenges in implementing the CEFR in English language classrooms. *LSP International Journal*, 8(2), 115–128. <https://doi.org/10.11113/lspi.v8.17977>
- Mohd Arof, K. Z., Ismail, S., & Saleh, A. L. (2018). Contractor's performance appraisal system in the Malaysian construction industry: Current practice, perception and understanding. *International Journal of Engineering & Technology*, 7(3.9), 46–53. <https://doi.org/10.14419/ijet.v7i3.9.15272>
- Moïse-Richard, A., Ménard, L., Bouchard, S., & Leclercq, A. L. (2021). Real and virtual classrooms can trigger the same levels of stuttering severity ratings and anxiety. *Journal of Fluency Disorders*, 68, 105830. <https://doi.org/10.1016/j.jfludis.2021.105830>
- Mukhtar, K., Javed, K., Arooj, M., & Sethi, A. (2020). Advantages, limitations and recommendations for online learning during COVID-19 pandemic era. *Pakistan Journal of Medical Sciences*, 36(COVID19-S4), S27–S31. <https://doi.org/10.12669/pjms.36.covid19-s4.2785>
- Mumba, B. (2024). Examining attitudes toward online learning classes amidst COVID-19 pandemic among students at the University of Lusaka. *Sage Open*, 14(2), 1–10. <https://doi.org/10.1177/21582440241236301>
- Musa, N. C., Nizam, A. S., Mohamad, W. N. W., & Othman, Z. (2021). Framing the English language (EL) CEFR-informed curriculum structure: The UKM experience. *Pertanika Journal of Social Sciences and Humanities*, 29(S3), 117–130. <https://doi.org/10.47836/pjssh.29.s3.13>
- Muttaqin, S., & Ida, I. (2015). Using English as a means of instruction in teaching content subjects. *Nobel Journal of Literature and Language Teaching*, 6(1), 47–60. <https://doi.org/10.15642/nobel.2015.6.1.47-60>

- Nabilah, J. (2021, May 27). *An investigation on perceived anxiety and speaking skills: The case of Indonesian EFL student teachers* [Paper presentation]. International Social Sciences and Innovation Congress, Ankara, Turkey. <https://doi.org/10.17605/OSF.IO/D75AM>
- Nawai, R., & Said, N. E. M. (2020). Implementation challenges of Common European Framework of Reference (CEFR) in a Malaysian setting: Insights on English teachers' attitude. *International Journal of Academic Research in Business and Social Sciences*, 10(7), 784–794. <https://doi.org/10.6007/ijarbss/v10-i7/7394>
- Nawawi, N. M., Zuhaimi, N., Sabu, K., Mahamud, N. S. R., & Nasir, N. A. M. (2021). CEFR for languages and its effective implementation in secondary schools in Malaysia. *Asian Journal of Assessment in Teaching and Learning*, 11(1), 63–72. <https://doi.org/10.37134/ajatel.vol11.1.6.2021>
- Ng, H., & Ahmad, M. (2021). Secondary school English teachers' knowledge and practice on CEFR-aligned English curriculum. *Asia Pacific Journal of Educators and Education*, 36(1), 75–91. <https://doi.org/10.21315/apjee2021.36.1.5>
- Nguyen Thi Dieu, P. (2024). *Investigating Vietnamese students' English language anxiety: An online TBLT course* [Doctoral thesis, Sheffield Hallam University]. Sheffield Hallam University Research Archive. https://shura.shu.ac.uk/34557/1/Phuong_2024_PhD_InvestigatingVietnameseStudents%27.pd
- Nirmalasari, N., Wahyudin, D., & Darmawan, D. (2024). Development of alternative thinking strategy curriculum design to improve academic, social and emotional competence. *International Journal of Learning, Teaching and Educational Research*, 23(3), 48–66. <https://doi.org/10.26803/ijlter.23.3.3>
- Nugraha, E. (2022). The CIPP model: Evaluation of the football school curriculum in Indonesia. *Jurnal Pendidikan Jasmani dan Olahraga*, 7(2), 262–267. <https://doi.org/10.17509/jpjo.v7i2.50459>
- Nurhayati, R., Hajaroh, M., Purwastuti, L. A., & Rohman, A. (2024). Training on education policy analysis for education stakeholders in Brebes. *JPPM (Jurnal Pendidikan dan Pemberdayaan Masyarakat)*, 11(1), 99–112. <https://doi.org/10.21831/jppm.v11i1.72653>
- Omid, A., Moghadas-Dastjerdi, T., & Yamani, N. (2020). Evaluation of health experts' education program for becoming multiprofessionals (family health caregiver) regarding health system transformation plan: An application of CIPP model. *Journal of Education and Health Promotion*, 9, 227. https://doi.org/10.4103/jehp.jehp_75_20
- Peter, L., & Asmawi, A. (2023). Online remote ESL education challenges, opportunities and readiness among high school students during school closure. *JOALL (Journal of Applied Linguistics and Literature)*, 8(2), 281–302. <https://doi.org/10.33369/joall.v8i2.26821>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Rahmanu, I. W. E. D., & Santosa, M. H. (2022). Exploring polytechnic learners' perception on multimodality in MALL in business administration instruction. *Dinamika Pendidikan*, 17(1), 12–24. <https://doi.org/10.15294/dp.v17i1.31578>
- Rahmi, R. S., & Murtafi'ah, B. (2022). Students' anxiety in speaking class: An investigation in online and offline learning system. *Eduvelop: Journal of English Education and Development*, 6(1), 66–78. <https://doi.org/10.31605/eduvelop.v6i1.1754>

Rajani Na Ayuthaya, J., & Sitthitikul, P. (2016). World Englishes-based lessons: Effects on anxiety and language achievement. *Arab World English Journal*, 7(2). <https://doi.org/10.24093/awej/vol7no2.14>

Ramulumo, M. (2023). Science students' attitudes towards the use of indigenous language in understanding visual representations. *Journal of Innovative Science Education*, 12(2), 125–140. <https://doi.org/10.15294/jise.v12i2.72094>

Ramzan, M., Javaid, Z. K., Kareem, A., & Mobeen, S. (2023). Amplifying classroom enjoyment and cultivating positive learning attitudes among ESL learners. *Pakistan Journal of Humanities and Social Sciences*, 11(2), 2298–2308. <https://doi.org/10.52131/pjhss.2023.1102.0522>

Razali, F., Sulaiman, T., & Ayub, A. F. M. (2022). Factors of learning towards creating blended learning curriculum using learning management system in higher education during COVID-19. *International Journal of Instruction*, 15(4), 723–744. <https://doi.org/10.29333/iji.2022.15439a>

Sahib, F. H. and Stapa, M. (2021). The impact of implementing the Common European Framework of Reference on language education: A critical review. *International Journal of Academic Research in Business and Social Sciences*, 11(11). <https://doi.org/10.6007/ijarbss/v11-i11/11160>

Samad, N. A. and Haron, Z. (2021). Formative assessment: Master trainers' attitude and practice. *The English Teacher*, 50(1), 1-13. <https://doi.org/10.52696/iwlm6113>

Savski, K. (2021). CEFR as language policy: Opportunimties and challenges for local agency in a global era. *The English Teacher*, 50(2), 60–70. <https://doi.org/10.52696/aide2513>

Shak, M. S. Y., Albakri, I. S. M. A., Tahir, M. H. M., & Adam, M. H. M. (2021). The use of imported CEFR-aligned English language textbooks in Malaysian schools: Issues and concerns. *International Journal of Academic Research in Business and Social Sciences*, 11(9). <https://doi.org/10.6007/ijarbss/v11-i9/10875>

Sharma, R. and Kumar, R. (2023). Increasing public participation to ensure holistic view in educational policies/action plans. *Indian Journal of Public Administration*, 69(3), 597-608. <https://doi.org/10.1177/00195561231177039>

Slykerman, R., Li, E., & Mitchell, E. (2022). Students' experience of online university education during the COVID-19 pandemic: Relationships to psychological health. *Student Success*, 13(1), 32–40. <https://doi.org/10.5204/ssj.2023>

Stognieva, O., & Popov, V. (2021). Assessment under COVID-19: Exploring undergraduate students' attitudes towards their online thesis proposal presentations vs. face-to-face. *Journal of Language and Education*, 7(4), 139–155. <https://doi.org/10.17323/jle.2021.11657>

Sudewi, P. W., & Fadilah, N. (2023). Students' perception of audio visual in teaching listening skills on the second grade students of SMP Negeri 3 Majene. *JET (Journal of English Teaching)*, 9(1), 98–107. <https://doi.org/10.33541/jet.v9i1.4578>

Suryana, D. (2023). CIPP evaluation model: Analysis of education implementation in PAUD based on government policy on implementation of learning during the COVID-19 pandemic. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 7(4), 4386–4396. <https://doi.org/10.31004/obsesi.v7i4.3722>

Tan, K. H., Jospa, M. E. A. W., Said, N. E. M., & Awang, M. M. (2021). Speak like a native English speaker or be judged: A scoping review. *International Journal of Environmental Research and Public Health*, 18(23), 12754. <https://doi.org/10.3390/ijerph182312754>

Tan, K. M., Foong, C. C., Adams, D., Hong, W., Lee, Y. K., & Pallath, V. (2023). Remote learning readiness amidst the COVID-19 pandemic amongst undergraduate medical students. *The Asia Pacific Scholar*, 8(4), 40–45. <https://doi.org/10.29060/taps.2023-8-4/sc3010>

Tan, T. T., Inciong, G. D., Ignacio, S. D., & Tiangco, A. C. S. (2022). Knowledge, skills, and attitudes of faculty members and residents-in-training of the Department of Rehabilitation Medicine of the Philippine General Hospital towards virtual learning: A cross-sectional study. *Acta Medica Philippina*, 56(4). <https://doi.org/10.47895/amp.v56i4.4219>

Teoh, G., & Liao, A. (2021). Emotional intelligence in distance learning: A case study of English as a second language via distance learning. *Journal of Language and Education*, 7(3), 151–165. <https://doi.org/10.17323/jle.2021.12624>

Tran, V. H., Wang, C., McLeod, S., & Verdon, S. (2021). Vietnamese–Australian children’s language proficiency and use. *International Journal of Bilingualism*, 25(6), 1735–1763. <https://doi.org/10.1177/13670069211034587>

Tsiplakides, I., & Keramida, A. (2009). Helping students overcome foreign language speaking anxiety in the English classroom. *International Education Studies*, 2(4). <https://doi.org/10.5539/ies.v2n4p39>

Tunç, F. (2010). *Evaluation of an English language teaching program at a public university using CIPP model* [Master’s thesis, Middle East Technical University]

Tuttle, B. E., & Munalim, L. O. (2022). Millennial students’ metalinguistic knowledge on headlines using grammaticality judgment test. *Studies in English Language and Education*, 9(3), 1093–1111. <https://doi.org/10.24815/siele.v9i3.23731>

Ulfatun, T., Septiyanti, F., & Lesmana, A. G. (2021). University students’ online learning self-efficacy and self-regulated learning during the COVID-19 pandemic. *International Journal of Information and Education Technology*, 11(12), 597–602. <https://doi.org/10.18178/ijiet.2021.11.12.1570>

Ulum, Ö. (2016). Evaluation of English as a foreign language program using CIPP model. *European Journal of English Language Teaching*. <https://doi.org/10.46827/ejel.v0i0.377>

Uri, N., & Aziz, M. (2018). Implementation of CEFR in Malaysia: Teachers’ awareness and the challenges. *3L: The Southeast Asian Journal of English Language Studies*, 24(3), 168–183. <https://doi.org/10.17576/3l-2018-2403-13>

Uri, N., & Aziz, M. (2020). The appropriacy and applicability of English assessment against CEFR global scale: Teachers’ judgement. *3L: The Southeast Asian Journal of English Language Studies*, 26(3), 53–65. <https://doi.org/10.17576/3l-2020-2603-05>

Wang, C. (2022). Comprehensively summarizing what distracts students from online learning: A literature review. *Human Behavior and Emerging Technologies*, 2022, Article 1483531. <https://doi.org/10.1155/2022/1483531>

Wang, P., Ganushchak, L., Welie, C., & van Steensel, R. (2025). Same anxiety, different faces. *Brain and Behavior*, 3(4), e70040. <https://doi.org/10.1002/brx2.70040>

Wang, X. (2022). Enhancing Chinese EFL students’ academic engagement: The impact of L2 enjoyment and academic motivation. *Frontiers in Psychology*, 13, Article 914682. <https://doi.org/10.3389/fpsyg.2022.914682>

Wati, S. (2023). Analyzing language attitude of Javanese and Sundanese urban community on English language pedagogy in Indonesia. *Qubahan Academic Journal*, 3(4), 298–305. <https://doi.org/10.58429/qaj.v3n4a182>

Wong, T., & Ismail, L. (2023). Relationship between language attitudes towards task-based language teaching and English language speaking performance among Malaysian polytechnic students. *AJELP: The Asian Journal of English Language and Pedagogy*, 11(2), 35–45. <https://doi.org/10.37134/ajelp.vol11.2.4.2023>

Yastibaş, A. E., & Kavgacı, T. (2020). Evaluating English for academic purposes II course through the CIPP model. *Gümüşhane University Journal of Social Sciences*, 11(1), 86–94. <https://doi.org/10.36362/gumus.615998>

Yi, W., & Wei, L. (2022). Multilingual learning and cognitive restructuring: The role of audiovisual media exposure in Cantonese–English–Japanese multilinguals' motion event cognition. *International Journal of Bilingualism*, 27(3), 331–348. <https://doi.org/10.1177/13670069221085565>

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, 11(2), 49–54. <https://doi.org/10.21315/eimj2019.11.2.6>

Zaki, A. (2021). The implementation of CEFR in ESL learning: Why does it matter to the Malaysian education system? *AJATeL*, 11(2), 1–13. <https://doi.org/10.37134/ajatel.vol11.2.1.2021>

Zaki, A. S., & Ahmad, A. (2017). The level of integration among students at secondary school: A study in Limbang, Sarawak. *The International Journal of Social Sciences and Humanities Invention*, 4(2). <https://doi.org/10.18535/ijsshi/v4i2.05>

Zaki, A. W. & Darmi, R. (2021). The implementation of CEFR in ESL learning: Why does it matter to the Malaysian education system?. *Asian Journal of Assessment in Teaching and Learning*, 11(2), 1-13. <https://doi.org/10.37134/ajatel.vol11.2.1.2021>

Zulkflee, Z., Marimuthu, P., & Tahir, M. (2023). Speaking anxiety during MUET–CEFR aligned oral presentations among ESL learners in Perak. *World Journal of English Language*, 13(2), 176–186. <https://doi.org/10.5430/wjel.v13n2p176>

Zulkiflee, M. F. A., & Nimehchisalem, V. (2022). Speaking anxiety among single-gender and co-educational schools' students in an online learning context. *Open Journal of Modern Linguistics*, 12(4), 438–459. <https://doi.org/10.4236/ojml.2022.124032>

Investigating Factors Influencing Personalized Mathematics Learning Adoption among Pre-university Students at the Maldives National University: An Extended UTAUT 2 Perspective

ABSTRACT

Authors: Mausooma Mohamed^{1,4}, Ahmad Fauzi Mohd Ayub^{2*}, and Nur Raihan Che Nawi³

Affiliation:¹Institute for Mathematical Research, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

²Department of Foundation of Education, Faculty of Educational Studies, University Putra Malaysia, 43400 Serdang, Malaysia.

³Department of Professional Development and Continuing Education, Faculty of Educational Studies, University Putra Malaysia, 43400 Serdang, Malaysia.

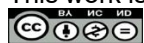
⁴Centre for Foundation Studies, The Maldives National University, Male' City, Maldives.

*Corresponding author:

afmy@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026|
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Learning Management Systems (LMS) are widely used to promote personalized learning in mathematics. However, sustained adoption remains uneven, particularly when students' study routines and prior experience with formal e-learning vary. This study investigates the determinants of pre-university students' behavioural intention (BI) to adopt Moodle-supported Personalized Mathematics Learning (PML) at The Maldives National University (MNU), where PML is implemented through structured Moodle-based resources, practice, and feedback, using an extended UTAUT2-informed framework that incorporates learning- and engagement-relevant constructs. A cross-sectional survey design was employed (N = 120; population frame = 559). A total of 47 items measured performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, learning value, habit, student commitment, intrinsic motivation, and behavioural intention. Measurement quality was assessed using exploratory factor analysis (KMO = 0.871; Bartlett's χ^2 (1081) = 4856.158, $p < .001$; 10-factor solution; 74.887% cumulative variance explained) and internal consistency reliability (Cronbach's α range = 0.744 - 0.924 across constructs). Pearson correlations indicated that BI was positively associated with all predictors, with the strongest bivariate relationships observed for habit and student commitment. In multiple regression with diagnostic checks, the model explained substantial variance in BI ($R^2 = 0.628$; adjusted $R^2 = 0.598$), with habit ($\beta = 0.484$, $p < .001$), learning value ($\beta = 0.220$, $p = .023$), and student commitment ($\beta = 0.261$, $p = .017$) emerging as significant unique predictors. The findings suggest that adoption in this context is best supported by implementation strategies that strengthen perceived learning value, sustain commitment, and foster routinised engagement that can evolve into habit. Results are interpreted as associations, rather than causal effects, and future work should incorporate behavioural usage indicators and longitudinal designs to connect intention to actual adoption and learning outcomes.

Keywords: Moodle, Personalized Mathematics Learning, Behavioral intention, Habit formation, UTAUT2, Learning Management System, Technology acceptance, Maldives

1. Introduction

Digital learning platforms are now routinely positioned as key enablers of personalised learning in mathematics, particularly where institutions seek to support differentiated instruction, continuous practice, and timely feedback at scale (Major et al., 2021; Shemshack & Spector, 2020). However, the effectiveness of such systems depends on more than availability: learners must be willing to adopt and sustain use (Abbad, 2021; Salloum & Shaalan, 2019). In practice, many LMS implementations achieve initial exposure without achieving durable uptake, especially in contexts where students' prior experience with structured e-learning varies and where learning routines are still developing (Akwene, 2024; Al-Azawei et al., 2016; Brown et al., 2022).

In the Maldives, the national push toward digitally supported learning has expanded the use of Learning Management Systems (LMS) across education levels (Azlifa & Saeed, 2021). At The Maldives National University (MNU), Moodle has been utilised to support Personalized Mathematics Learning (PML) by delivering resources, structured practice, and interactive learning activities (Mohamed, 2020). In this study, Moodle-supported PML refers to the use of Moodle to organise mathematics learning through structured learning resources, repeated practice opportunities, interactive activities, and feedback processes intended to support sustained and differentiated student engagement in mathematics. Accordingly, the term does not refer to generic Moodle use in any course, but specifically to Moodle-supported mathematics learning as experienced by the pre-university students examined in this study. While Moodle provides a technically capable environment, the success of Moodle-supported PML depends on students' behavioural intention to engage with the platform consistently (Abbad, 2021; Raza et al., 2021). Understanding what most strongly aligns with students' intention is therefore practically important for instructional design, institutional planning, and sustainable LMS-supported innovation in mathematics education (Gamage et al., 2022; Ismail et al., 2021).

Research on technology adoption has frequently drawn on the Unified Theory of Acceptance and Use of Technology (UTAUT) and its extension, UTAUT2, which explain intention through constructs such as performance expectancy, effort expectancy, social influence, facilitating conditions, and habit (Venkatesh et al., 2003; Venkatesh et al., 2012; Venkatesh et al., 2016). However, evidence from educational settings has been mixed: predictors that perform strongly in workplace or consumer contexts do not always show the same explanatory pattern among students, particularly when learning value perceptions, motivational processes, and persistence mechanisms are central to engagement (Abbad, 2021; Bervell & Umar, 2017; Dwivedi et al., 2019).

In Moodle-supported learning environments, students may not adopt primarily because an LMS is "easy" or "useful" in general terms, but because the learning experience becomes valuable, personally meaningful, and integrated into their study routine (Ain et al., 2016; Ortiz de Guinea & Markus, 2009; Zacharis & Nikolopoulou, 2022). This study is positioned in relation to recent educational extensions of UTAUT2, particularly Ain et al. (2016), who foreground learning value, and Zacharis and Nikolopoulou (2022), who applied an extended UTAUT2 approach to e-learning platform adoption. What is distinctive, is the focus on Moodle-supported mathematics learning in the Maldivian pre-university context and the simultaneous examination of learning value, student commitment, and intrinsic motivation alongside the canonical UTAUT2 predictors to identify implementation-relevant priorities. To address this knowledge gap, this study examines pre-university students' behavioural intention to adopt Moodle-supported PML using an extended UTAUT2 model (Venkatesh et al., 2012; Zacharis & Nikolopoulou, 2022). In addition to the canonical predictors, the study foregrounds education-related drivers of engagement, especially learning value, student commitment, and intrinsic motivation, which can distinguish intention in a learning environment where sustained engagement is needed (Ain et al., 2016; San-Martín et al., 2020; Wilkins et al., 2016). The study provides empirical data on a setting underrepresented in the overall literature on LMS adoption by focusing on pre-university students in the Maldives (Adam, 2017; Mumthaz, 2021).

This study aims to identify which determinants are most closely associated with pre-university students' behavioural intention to adopt Moodle-supported PML. The study has three contributions: (1) a contextual contribution by providing evidence from the Maldives pre-university context, where the adoption of Moodle-supported PML is institutionally relevant (Adam, 2017; Azlifa & Saeed, 2021); (2) an applied contribution by determining which levers are most aligned with intention in this setting, which supports implementation priorities **for** administrators and instructors (Akwene, 2024; Brown et al., 2022); and (3) a theoretical contribution through refining a UTAUT2-based framework by incorporating value- and persistence-related constructs relevant to sustained learning engagement (Ain et al., 2016; San-Martín et al., 2020; Venkatesh et al., 2012; Zacharis & Nikolopoulou, 2022). Accordingly, the study addresses the following research questions: What are the levels of pre-university students' perceptions of Moodle-supported PML across the study constructs?

RQ1: What are the bivariate relationships between behavioural intention (BI) and each predictor construct?

RQ2: Which constructs uniquely predict behavioural intention (BI) when all predictors are modelled simultaneously?

2. Literature Review and Conceptual Framework

Personalized learning in mathematics is often implemented through structured cycles of practice, feedback, and adaptive support, in which learners engage with content at an appropriate pace and level of difficulty (Major et al., 2021; Shemshack & Spector, 2020). Learning Management Systems (LMS) such as Moodle can operationalize these cycles by hosting learning resources, sequencing tasks, enabling formative assessment, and providing feedback loops that sustain engagement (Gamage et al., 2022; Ismail et al., 2021; Molins & García, 2023). Nevertheless, personalization facilitated by LMSs is not implemented automatically because the platform is there; instead, its adoption requires students to feel that the learning process is a worthy, manageable, and stable one, aligned with their study habits (Ain et al., 2016; Moorthy et al., 2019; Zacharis & Nikolopoulou, 2022). For this reason, adoption research in education frequently draws on theories of technology acceptance and use to explain why learners intend to engage with digital learning platforms (Abbad, 2021; Salloum & Shaalan, 2019).

A widely used theoretical lens is the Unified Theory of Acceptance and Use of Technology (UTAUT), which explains behavioural intention through performance expectancy, effort expectancy, and social influence, while acknowledging enabling conditions as drivers of use (Venkatesh et al., 2003). UTAUT2 extends this model by incorporating additional determinants, particularly hedonic motivation and habit, to improve explanatory power in contexts where users have discretion in adoption decisions (Venkatesh et al., 2012). The UTAUT2 model, based on Venkatesh et al. (2022) is presented in Figure 1.

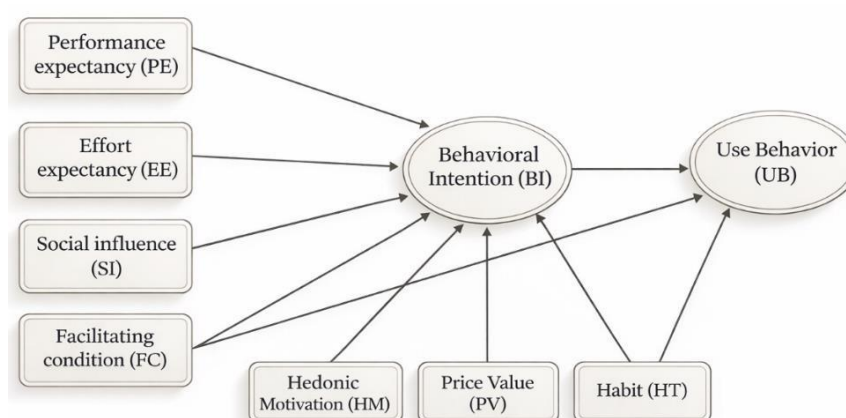


Figure 1. UTAUT2 Model (author's redrawing based on Venkatesh et al., 2012).

In educational settings, UTAUT2 has been applied to explain students' intention to use LMS platforms, digital learning tools, and online learning resources, largely because it offers a structured account of both instrumental beliefs (usefulness and ease) and experiential drivers (enjoyment and habit) (Abbad, 2021; Azizi et al., 2020; Raza et al., 2021; Salloum & Shaalan, 2019). Nevertheless, findings in education contexts have often been mixed: predictors that are robust in workplace or consumer settings do not always emerge as uniquely significant for students once multiple factors are modeled simultaneously (Abbad, 2021; Bervell & Umar, 2017; Zacharis & Nikolopoulou, 2022). This inconsistency is plausible because learning-related technology adoption is not only a "tool use" decision; it is embedded in self-regulation, persistence, and perceived learning payoff, which can overlap with or overshadow classic expectancy-based predictors when they are modeled together (Ain et al., 2016; Ortiz de Guinea & Markus, 2009; San-Martín et al., 2020). More specifically, prior studies suggest that expectancy- and support-related constructs such as performance expectancy, effort expectancy, social influence, and facilitating conditions do not always retain stable unique effects in educational LMS models once motivational and engagement-related predictors are considered, whereas habit, learning value, and context-specific persistence-related factors often remain more salient for continuance-oriented outcomes (Abbad, 2021; Ain et al., 2016; Bervell & Umar, 2017; Zacharis & Nikolopoulou, 2022). This pattern provides an empirical basis for extending UTAUT2 in learning instead of relying only on the canonical predictor set.

Against this background, the present study adopts a UTAUT2-informed foundation and extends it with education-relevant engagement constructs that are especially salient for sustained participation in Moodle-supported Personalized Mathematics Learning (PML) (Ain et al., 2016; Venkatesh et al., 2012; Zacharis & Nikolopoulou, 2022). Specifically, the model includes performance expectancy (PE), effort expectancy (EE), social influence (SI), facilitating conditions (FC), hedonic motivation (HM), and habit (HT), and extends the framework by foregrounding learning value (LV), student commitment (SC), and intrinsic motivation (IM) as additional explanatory drivers of behavioural intention (BI) (San-Martín et al., 2020; Venkatesh et al., 2012; Wilkins et al., 2016; Zacharis & Nikolopoulou, 2022). The inclusion of these three constructs is deliberate. Learning value is retained as prior educational extensions have shown that students' judgments about the worth and utility of learning through an LMS can contribute beyond general usefulness perceptions (Ain et al., 2016; Zacharis & Nikolopoulou, 2022). Student commitment is included as mathematics engagement often depends on persistence, effort investment, and willingness to continue despite delayed rewards (San-Martín et al., 2020; Wilkins et al., 2016). While, intrinsic motivation considered as autonomous interest in learning activities may support voluntary participation in Moodle-supported practice even when external pressure is limited (Alamri et al., 2020; Deci & Ryan, 1985; Oudeyer & Kaplan, 2009).

The conceptual logic is that students are more likely to intend to adopt Moodle-supported PML when they expect the platform to support learning performance (PE), perceive it as manageable to use (EE), experience normative encouragement (SI), perceive adequate support and resources (FC), experience enjoyment during use (HM), perceive strong learning value from engagement (LV), possess commitment to persist (SC), experience intrinsic interest in learning activities (IM), and develop habitual engagement patterns that stabilize continued intention (HT) (Ortiz de Guinea & Markus, 2009; Tamilmani et al., 2019; Venkatesh et al., 2012). Performance expectancy refers to the extent to which students believe that using Moodle-supported PML will enhance their mathematics learning effectiveness or outcomes (Venkatesh et al., 2003; Venkatesh et al., 2012). Effort expectancy reflects perceived ease of use and the cognitive effort required to engage with the system (Venkatesh et al., 2003; Venkatesh et al., 2012). Although PE and EE often co-occur empirically, they are conceptually distinct: PE concerns the expected academic benefit of using Moodle-supported PML, whereas EE concerns the ease and mental effort involved in using it. A platform may be easy to use without being perceived as educationally valuable, and it may be educationally valuable while still requiring substantial effort. For this reason, both constructs are retained, while acknowledging that they may not always remain strongly discriminant in educational contexts. Social influence reflects perceived encouragement from instructors, peers, or the institution to use Moodle-supported PML (Venkatesh et

al., 2003; Venkatesh et al., 2012; Raza et al., 2021). Facilitating conditions capture perceived access to necessary resources and support (such as reliable connectivity, access to devices, and guidance), enabling students to engage with Moodle consistently (Al-Azawei et al., 2016; Venkatesh et al., 2012). Hedonic motivation refers to the degree to which students experience enjoyment or positive affect while using Moodle-supported PML, which may strengthen their willingness to continue engaging (Tamilmani et al., 2019; Venkatesh et al., 2012). Habit reflects automaticity and routinized behaviour; as Moodle engagement becomes part of students' regular study pattern, intention becomes less dependent on deliberation and more dependent on continued routine (Moorthy et al., 2019; Ortiz de Guinea & Markus, 2009; Venkatesh et al., 2012).

Because Moodle-supported PML is also a learning practice that requires persistence over time, the model emphasises three education-relevant constructs. Learning value refers to students' appraisal of the usefulness and worthwhileness of the learning experience itself, whether Moodle-supported PML is perceived as beneficial, meaningful, and aligned with learning goals (Ain et al., 2016; Zacharis & Nikolopoulou, 2022). Student commitment captures the learner's willingness to persist, invest effort, and maintain engagement over time, particularly when tasks require repeated practice and delayed rewards (San-Martín et al., 2020; Wilkins et al., 2016). Intrinsic motivation refers to engagement driven by inherent interest or satisfaction in learning activities, consistent with motivational perspectives that emphasise autonomous interest as a driver of sustained learning engagement (Alamri et al., 2020; Deci & Ryan, 1985; Oudeyer & Kaplan, 2009). In Moodle-supported PML, intrinsic motivation can support voluntary engagement with practice activities and reinforce persistence when external pressures are minimal (Alamri et al., 2020). To reduce conceptual ambiguity, the three constructs are treated as related but non-equivalent. Learning value represents a cognitive appraisal that Moodle-supported PML is worthwhile for learning; student commitment represents persistence and sustained effort toward continued participation; and intrinsic motivation represents inherent interest and satisfaction in the learning activity itself. This distinction is also important for separating intrinsic motivation from hedonic motivation: HM refers primarily to enjoyment associated with using the platform, whereas IM refers to inherent interest in engaging with the mathematics learning activity supported by the platform. The two may overlap, but they are not conceptually identical, because enjoyment of system use does not necessarily imply autonomous interest in learning, and autonomous interest in learning does not necessarily depend on the platform being enjoyable in a hedonic sense.

The present study focuses on the main effects of these predictors on behavioural intention. Although UTAUT2 includes moderators such as age, gender, and experience, moderation testing typically requires larger sample sizes to preserve statistical power and to stabilize the estimation of interaction effects (Venkatesh et al., 2012). Given the study's sample size and its objective of identifying salient direct predictors of intention for applied implementation, moderators are not included in the model. They are prioritized for future work, particularly in follow-up designs that can incorporate prior LMS exposure and other contextual factors likely to shape transferability (Bervell & Umar, 2017; Dwivedi et al., 2019). The hypothesised model for the present study, adapted from UTAUT2 and extended with learning value, student commitment, and intrinsic motivation, is presented in Figure 2.

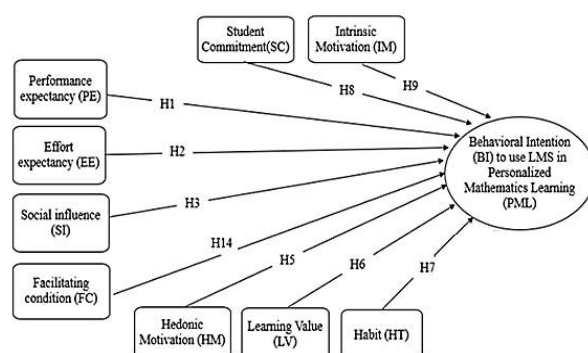


Figure 2. Hypothesised model for Moodle-supported PML adoption (adapted from Venkatesh et al., 2012).

Based on the conceptual framework, the study proposes the following hypotheses:

- H_{a1}: Performance expectancy (PE) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a2}: Effort expectancy (EE) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a3}: Social influence (SI) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a4}: Facilitating conditions (FC) have a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a5}: Hedonic motivation (HM) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a6}: Learning value (LV) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a7}: Habit (HT) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a8}: Student commitment (SC) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.
- H_{a9}: Intrinsic motivation (IM) has a positive and significant effect on behavioural intention (BI) to adopt Moodle-supported PML.

3. Methodology

3.1. Research Design and Setting

The study employed a quantitative, cross-sectional correlational research design to identify variables related to behavioural intention (BI) of pre-university students to use Moodle-based Personalised Mathematics Learning (PML) at The Maldives National University (MNU) (Pratama et al., 2023). Moodle is the institution's Learning Management System (LMS) used to facilitate the delivery of PML, learning resources, and student learning activities (Mohamed, 2020). For the purposes of this study, Moodle-supported PML refers to the use of Moodle to deliver structured mathematics learning resources, repeated practice opportunities, and feedback-oriented learning activities intended to support sustained and differentiated engagement in mathematics learning.

3.2. Participants, Sampling, and Data Collection

The study population comprised 559 pre-university students at the Maldives National University (MNU). All 559 eligible students were invited to participate via an online questionnaire administered using Google Forms, and recruitment was closed once the target sample size was achieved. A total of 120 students completed the survey, yielding a response rate of 21.5% (120/559). The inclusion criterion was enrolment as a pre-university student at MNU; no exclusion criteria were applied. Because participation was voluntary and no random selection procedure was used, the achieved sample is best interpreted as a self-selected, non-probability sample rather than as a statistically representative sample of the full invited population. Convenience sampling was adopted to allow efficient access to available students; however, this approach limits generalizability beyond the sampled cohort (Akwene, 2024). The survey was administered during the study period under institutional ethical approval (Approval No.: RR/2023/18).

3.3. Instrument Development and Measures

Data were collected using a structured questionnaire comprising 47 items measuring 10 constructs: PE, EE, SI, FC, HM, LV, HT, SC, IM, and BI. Items were rated on a five-point Likert scale ranging from Strongly disagree (1) to Strongly agree (5). The instrument was adapted from previously validated scales and refined for the PML/Moodle context (Ain et al., 2016; Venkatesh et al., 2012; Zacharis & Nikolopoulou, 2022). Content suitability and clarity were reviewed by five experts before administration. The final item wording and scale anchors are provided in Appendix A. Appendix A also reports the final retained items by construct and documents any items removed during post-administration measurement review.

3.4. Construct Validity and Internal Consistency

Construct validity was assessed using a staged measurement evaluation procedure. First, Exploratory Factor Analysis (EFA) was used to examine the factor structure of the predictor-item pool (PE, EE, SI, FC, HM, LV, HT, SC, and IM). To avoid circularity between predictor validation and outcome measurement, BI items were evaluated separately rather than being included in the same EFA as the predictor constructs. The suitability of the data for factor analysis was evaluated using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett’s test of sphericity. Factor extraction used Principal Axis Factoring, and rotation used Direct Oblimin. Because the constructs in this study were theoretically expected to correlate, an oblique rotation was preferred.

Factor retention and item adequacy were evaluated using factor loadings and cross-loadings; items with weak loadings or problematic cross-loadings were reviewed for potential removal. Items demonstrating weak primary loadings, problematic cross-loadings, or poor corrected item-total behaviour were reviewed for removal, and all retained and removed items are documented in Appendix A. Internal consistency reliability was examined using Cronbach’s alpha for each construct. Where item removal improved reliability and measurement coherence, items were removed, and the revised reliability results were retained. The final reliability coefficients (and retained item counts) for each construct are reported in the Findings section, with the full measurement outputs available in the supplementary materials.

Given the size of the achieved sample ($N = 120$) relative to the original item pool, the EFA results are interpreted cautiously as an initial assessment of measurement structure rather than as definitive evidence of a fully stable factor solution. Discriminant validity was additionally evaluated using average variance extracted (AVE) and the Fornell–Larcker criterion for the final construct set.

3.5. Data Analysis Strategy and Diagnostic Checks

Data were analysed using IBM SPSS Statistics (Version 25). Descriptive statistics (means, standard deviations, skewness, and kurtosis) were computed to summarize construct distributions. Pearson correlation analysis was used to examine bivariate associations between behavioural intention (BI) and predictor constructs.

To test the multivariate model, multiple linear regression was estimated with BI as the dependent variable and PE, EE, SI, FC, HM, LV, HT, SC, and IM as predictors. Regression assumptions were evaluated and reported, including multicollinearity (Variance Inflation Factor and tolerance), independence of errors (Durbin–Watson), residual normality (histogram and P–P plot), and homoscedasticity/linearity (standardized residuals versus standardized predicted values). Collinearity diagnostics were further examined using condition indices and variance proportions as reported in the SPSS collinearity diagnostics output.

Given the number of predictors relative to the sample size ($N = 120$; nine predictors), the analysis is interpreted as identifying associations and estimating the relative unique contribution of predictors rather than as a definitive test of small effects (Pratama et al., 2023). While common rules of thumb (e.g., 10–15 cases per predictor) suggest the sample is acceptable for estimation, statistical power is

limited for detecting small incremental effects and interaction terms. For this reason, the study focuses on main effects and prioritizes transparent reporting of coefficients, confidence intervals, and diagnostics.

Bivariate correlations are reported to describe relationships among constructs and to contextualize the multivariate findings. The regression model is treated as the primary inferential analysis, while correlations are interpreted as supportive rather than as independent hypothesis tests; accordingly, emphasis is placed on the consistency of evidence across regression and correlation results rather than on isolated p-values.

Although UTAUT2 includes canonical moderators (e.g., age, gender, experience, voluntariness), these were not modelled as interaction terms in the present study because moderation testing typically requires larger samples to support stable interaction estimation and adequate power (Venkatesh et al., 2012). Moderators are therefore positioned as a priority for future research, particularly in designs that can incorporate prior LMS exposure and other contextual constraints relevant to Moodle-supported learning (Bervell & Umar, 2017).

Lastly, robustness was addressed through diagnostic review and transparent reporting of findings, with a specific focus on the possibility of shared variance among predictors. To examine coefficient stability under correlated-predictor conditions, supplementary sensitivity analyses were performed using hierarchical regression. Still, the evidence most commonly used in this study is based on the reported OLS model and its diagnostic tests.

3.6. Ethical Considerations

Ethical approval was obtained from the Research Development Office (RDO) ethics committee at MNU (Approval No.: RR/2023/18). Participation was voluntary, and responses were anonymized for analysis. Data were stored securely and used exclusively for research purposes.

3.7. Data/Materials Availability

The questionnaire instrument (item wording and anchors) is provided in Appendix A. The de-identified dataset is deposited in an appropriate repository and is referenced in the Data Availability statement in this manuscript.

4. Findings

4.1. Measurement Quality: Construct Validity and Internal Consistency

Construct validity was assessed using exploratory factor analysis (EFA) of the predictor constructs, with behavioural intention (BI) evaluated separately from the predictor-item pool to avoid circularity between predictor validation and outcome measurement. Sampling adequacy was strong ($KMO = 0.871$), and Bartlett's test of sphericity was significant ($\chi^2 (1081) = 4856.158, p < .001$), indicating that the correlation matrix was suitable for factor analysis. Using Principal Axis Factoring with Direct Oblimin, a 10-factor solution consistent with the conceptual model was retained. The 10 factors explained 74.887% of the cumulative variance. Item performance was acceptable overall, with rotated factor loadings ranging from 0.420 to 0.824 and communalities ranging from 0.580 to 0.854.

Internal consistency (reliability) was assessed using Cronbach's alpha for each construct. Items demonstrating weak or negative corrected item-total behaviour (typically negatively keyed items) were removed to improve internal consistency. The final retained items and the item-retention log are

reported in Appendix A. The final reliability coefficients and retained item counts are reported in Table 2.

4.2. Descriptive Statistics

Construct-level descriptive statistics ($N = 120$) are presented in Table 1. Behavioural intention (BI) was generally high ($M = 3.948$, $SD = 0.762$). Facilitating conditions (FC) and student commitment (SC) recorded the highest mean levels, indicating that students generally perceived strong support and commitment regarding Moodle-supported PML. By contrast, social influence (SI) had the lowest mean ($M = 3.077$, $SD = 0.997$), suggesting that relative to other constructs, perceived peer, instructor, or institutional encouragement was less strongly endorsed in this sample. This pattern is descriptively noteworthy because it suggests that students' intentions may be shaped more strongly by individual engagement-related perceptions than by external social pressure. In addition, FC showed the highest mean ($M = 4.121$) and a relatively high kurtosis value (3.911), indicating a possible ceiling effect and restricted variability in perceived support conditions. This restricted range may help explain why FC correlates positively with BI but does not emerge as a unique predictor in the full multivariate model.

Table 1. Descriptive Statistics For Study Constructs ($N = 120$)

Construct	N	Mean	SD	Skewness	Kurtosis
PE	120	3.971	0.783	-0.253	-0.665
EE	120	3.921	0.821	-0.787	1.035
SI	120	3.077	0.997	0.127	-0.543
FC	120	4.121	0.636	-1.124	3.911
HM	120	3.590	0.902	-0.580	0.454
LV	120	3.731	0.787	-1.055	1.910
HT	120	3.583	0.869	-0.523	0.322
SC	120	4.005	0.751	-0.893	1.565
IM	120	3.915	0.828	-1.069	2.137
BI	120	3.948	0.762	-0.878	2.026

Note. Values are based on final retained items for each construct.

4.3. Reliability

Final internal consistency estimates met acceptable thresholds across constructs. The strongest reliabilities were observed for HM, PE, EE, and BI, while FC remained acceptable. Full reliability results are reported in Table 2.

Table 2. Internal Consistency Reliability by Construct (Final Retained Items)

Construct	Items retained (k)	Cronbach's α
PE	4	0.911
EE	4	0.908
SI	4	0.817
FC	4	0.744
HM	4	0.924
LV	4	0.831
HT	3	0.874
SC	5	0.849
IM	4	0.867
BI	4	0.900

4.4. Correlation Analysis (Pearson Correlations)

Pearson correlations among constructs are reported in Table 3. Behavioural intention (BI) showed the strongest positive associations with habit (HT) and student commitment (SC), followed by learning value (LV), intrinsic motivation (IM), and hedonic motivation (HM). These correlations support the expectation that commitment, habitual engagement, and motivational/value beliefs align with intention to adopt Moodle-supported PML. At the same time, several inter-predictor correlations were relatively high, notably PE–EE ($r = 0.798$), HM–HT ($r = 0.757$), HM–LV ($r = 0.728$), and SC–IM ($r = 0.713$), indicating substantial shared variance among conceptually related constructs. These patterns are important for interpreting the multivariate regression results, particularly where bivariate associations do not translate into unique regression effects.

Table 3. Pearson Correlation Matrix (Construct Level, N = 120)

	PE	EE	SI	FC	HM	LV	HT	SC	IM	BI
PE	—									
EE	0.798***	—								
SI	0.355***	0.345***	—							
FC	0.312***	0.387***	0.113	—						
HM	0.685***	0.624***	0.526***	0.387***	—					
LV	0.561***	0.488***	0.501***	0.477***	0.728***	—				
HT	0.508***	0.475***	0.441***	0.352***	0.757***	0.618***	—			
SC	0.627***	0.556***	0.321***	0.461***	0.713***	0.617***	0.709***	—		
IM	0.533***	0.479***	0.325***	0.445***	0.615***	0.593***	0.530***	0.713***	—	
BI	0.461***	0.413***	0.244**	0.460***	0.569***	0.600***	0.704***	0.694***	0.574***	—

Note. Lower triangle reported. Two-tailed tests. *** $p < .001$, ** $p < .01$, * $p < .05$.

4.5. Multiple Regressions Predicting Behavioural Intention (BI) + Diagnostics

Multiple linear regression was estimated with BI as the dependent variable and PE, EE, SI, FC, HM, LV, HT, SC, and IM as predictors. The model was statistically significant and explained substantial variance in BI ($R = 0.792$, $R^2 = 0.628$, Adjusted $R^2 = 0.598$, $F(9, 110) = 20.635$, $p < .001$).

As shown in Table 4, habit (HT) was the strongest unique predictor of BI (standardized $\beta = 0.484$, $p < .001$), followed by learning value (LV) ($\beta = 0.220$, $p = .023$) and student commitment (SC) ($\beta = 0.261$, $p = .017$). Other predictors were not statistically significant after controlling for the full set of constructs. Accordingly, Ha6, Ha7, and Ha8 were supported, whereas Ha1, Ha2, Ha3, Ha4, Ha5, and Ha9 were not supported in the multivariate model.

Assumption/diagnostic checks. Multicollinearity was within acceptable bounds (VIF range 1.53–4.23; see Table 4). Independence of errors was supported (Durbin–Watson = 1.833). Residual diagnostics (normality and homoscedasticity checks) did not indicate violations severe enough to invalidate inference for the present model.

Table 4. Multiple Regression Predicting BI (N = 120)

Predictor	B	SE(B)	β	t	p	95% CI for B	VIF
PE	0.033	0.107	0.034	0.313	.755	[-0.179, 0.245]	3.57
EE	-0.024	0.094	-0.026	-0.255	.799	[-0.210, 0.162]	3.02
SI	-0.082	0.055	-0.107	-1.488	.139	[-0.191, 0.027]	1.54
FC	0.127	0.086	0.106	1.477	.143	[-0.043, 0.298]	1.53
HM	-0.169	0.101	-0.201	-1.677	.096	[-0.370, 0.031]	4.23
LV	0.213	0.092	0.220	2.311	.023	[0.030, 0.396]	2.68
HT	0.424	0.086	0.484	4.935	<.001	[0.254, 0.595]	2.84
SC	0.265	0.109	0.261	2.423	.017	[0.048, 0.481]	3.43

IM	0.097	0.081	0.106	1.204	.231	[-0.063, 0.258]	2.28

Note. DV = BI. Unstandardized coefficients (B), standard errors, standardized coefficients (β), and 95% CIs are reported. VIF = variance inflation factor.

To facilitate interpretation, Table 5 summarizes the hypothesis-testing outcomes based on the multivariate regression model.

Table 5. Hypothesis Testing Summary

Hypothesis	Path	β	p	Decision
Ha1	PE $\rightarrow$ BI	0.034	.755	Not supported
Ha2	EE $\rightarrow$ BI	-0.026	.799	Not supported
Ha3	SI $\rightarrow$ BI	-0.107	.139	Not supported
Ha4	FC $\rightarrow$ BI	0.106	.143	Not supported
Ha5	HM $\rightarrow$ BI	-0.201	.096	Not supported
Ha6	LV $\rightarrow$ BI	0.220	.023	Supported
Ha7	HT $\rightarrow$ BI	0.484	<.001	Supported
Ha8	SC $\rightarrow$ BI	0.261	.017	Supported
Ha9	IM $\rightarrow$ BI	0.106	.231	Not supported

5. Discussion

5.1. Interpreting the Dominant Predictors: Habit, Learning Value, and Student Commitment

The most important interpretive result of this study is that behavioural intention to adopt Moodle-supported PML was explained primarily by habit, learning value, and student commitment rather than by the full canonical set of UTAUT2 predictors. The multivariate model explained substantial variance in BI ($R^2 = 0.628$; Adjusted $R^2 = 0.598$), indicating that the extended UTAUT2-based model provided a strong account of intention in this setting (Venkatesh et al., 2012; Zacharis & Nikolopoulou, 2022). This level of explanatory power is strong in relation to comparable educational UTAUT2 studies, although it is not unusually high. For example, El-Masri and Tarhini (2017) reported that their extended UTAUT2 e-learning model explained 68% of the variance in behavioural intention in the Qatari sample and 63% in the USA sample. In addition, recent synthesis work has noted that UTAUT2-based models can explain up to 74% of the variance in behavioural intention in technology-use contexts (Venkatesh et al., 2012; Zheng et al., 2025). Against this benchmark, the present model's R^2 of 0.628 indicates solid explanatory performance in this pre-university mathematics setting, while also suggesting that some relevant contextual or motivational influences remain outside the current model. Habit (HT) emerged as the strongest unique predictor, followed by student commitment (SC) and learning value (LV), suggesting that intention in this context is shaped most strongly by routinised engagement, persistence, and perceived worthwhileness of the learning experience. This pattern is also consistent with broader educational UTAUT2 evidence. Zheng et al. (2025), in a meta-analytic review of 91 higher-education e-learning studies involving 37,910 participants, found that habit was the most influential antecedent of behavioural intention toward e-learning. The same review also notes that Zacharis and Nikolopoulou (2022) identified habit as the strongest predictor in their university e-learning study.

Habit being the strongest predictor aligns with evidence in technology adoption showing that repeated use reduces deliberative effort and supports continuance intentions over time (Ortiz de Guinea & Markus, 2009; Venkatesh et al., 2012; Moorthy et al., 2019). In the present study, this suggests that once Moodle-supported mathematics learning becomes embedded in students' regular study routines, continued intention depends less on one-time judgments about the system and more on the stability of

those routines. This is especially relevant in pre-university settings, where study patterns may still be forming and where repeated exposure to structured digital tasks can gradually normalize platform use.

Student commitment and learning value also emerged as significant unique predictors of BI (Table 4). Together, these results point to an adoption mechanism that is both self-regulatory and instrumental (San-Martin et al., 2020; Wilkins et al., 2016; Ain et al., 2016; Zacharis & Nikolopoulou, 2022). Students appear more likely to intend continued use when they perceive Moodle-supported PML as genuinely beneficial for mathematics learning and when they are psychologically prepared to persist with the required engagement. This pattern is theoretically important because it suggests that in learning-specific technology contexts, value and persistence constructs may operate as more proximal drivers of intention than the more general expectancy constructs that dominate the original UTAUT2 formulation.

The differential pattern across the three extension constructs is also informative. Learning value and student commitment were significant, whereas intrinsic motivation was not. This suggests that not all education-specific extensions add explanatory power equally. In this sample, intention appears to depend more on whether students judge the learning process to be worthwhile and are willing to persist with it than on whether they experience inherent interest in the learning activity itself after other related predictors are controlled. Accordingly, future educational adaptations of UTAUT2 may benefit from prioritizing value- and persistence-related constructs over broader motivational constructs unless the latter are shown to contribute distinct variance in the specific learning context under study.

5.2. Why The Canonical UTAUT2 Predictors Were Not Uniquely Significant

Several UTAUT2 constructs showed meaningful bivariate associations with BI (Table 3), but none remained significant in the multivariate model (Table 4). This pattern is common when predictors share variance and suggests that their influence may be indirect or subsumed by more proximal determinants (Dwivedi et al., 2019; Bervell & Umar, 2017). The present findings therefore do not suggest that performance expectancy, effort expectancy, social influence, facilitating conditions, and intrinsic motivation are unimportant in absolute terms; rather, they suggest that their explanatory role was not unique once learning value, student commitment, and habit were entered simultaneously.

Performance expectancy (PE) and effort expectancy (EE) may have functioned as necessary but not differentiating conditions in this setting. Students reported relatively high mean levels on these expectancy constructs (Table 1), which may have reduced their ability to distinguish between students with stronger and weaker adoption intentions once more engagement-proximal variables were controlled. This interpretation is consistent with observations that expectancy constructs in educational LMS research may matter most at the stage of initial acceptance, while continuance-oriented outcomes depend more heavily on value, commitment, and routine (Abbad, 2021; Zacharis & Nikolopoulou, 2022).

Facilitating conditions (FC) also correlated positively with BI (Table 3) but did not retain unique significance in the full model (Table 4). This result is possibly linked to the descriptive pattern observed in Table 1: FC had the highest mean ($M = 4.121$) and a high kurtosis value (3.911), indicating a possible ceiling effect and restricted variability. When support conditions are already broadly perceived as adequate, they may cease to function as the primary differentiator of intention, leaving more variance to be explained by students' engagement-related judgments and routines (Akwene, 2024; Al-Azawei et al., 2016).

Social influence (SI) showed the lowest mean of all constructs ($M = 3.077$) and only a modest association with BI. This is substantively important for institutional interpretation. In this cohort, adoption appears to be less strongly driven by peer or authority encouragement than by students' own developing routines, perceived learning value, and persistence. In practical terms, this suggests that in this setting, institutional messaging alone is unlikely to be sufficient; social encouragement may help at

the margins, but it is unlikely to substitute for instructional designs that make Moodle-supported PML useful, repeatable, and worth sustaining. This finding may also reflect the individualized nature of mathematics practice, where continued engagement depends less on social visibility and more on task integration into students' private study behaviour.

Intrinsic motivation (IM) correlated positively with BI (Table 3) but became non-significant in the multivariate model (Table 4), implying that its explanatory role overlapped with learning value, commitment, and hedonic experience rather than remaining distinct after adjustment. This is consistent with the broader theoretical implication noted above: in this sample, sustained intention appears to be explained more directly by "this helps me learn" and "I will persist with this" than by "I am inherently interested in this activity" once the shared variance across motivational constructs is accounted for.

These results suggest that in this cohort, the canonical UTAUT2 constructs functioned more as background conditions than as decisive differentiators of intention. This does not disconfirm UTAUT2; rather, it indicates that when the model is applied to a learning-specific context that requires sustained effort, theory may need to foreground engagement continuity, perceived learning payoff, and persistence more explicitly.

5.3. Reconciling Hedonic Motivation: Positive Bivariate Association but Negative Multivariate Coefficient

A key concern is the apparent contradiction that hedonic motivation (HM) correlated positively with BI (Table 3) but showed a negative, non-significant coefficient in the multivariate model (Table 4). This is not necessarily an error; it is consistent with shared variance or suppression when correlated motivational predictors are entered simultaneously (Dwivedi et al., 2019; Tamilmani et al., 2019).

HM overlapped strongly with other engagement-related predictors, particularly HT, SC, and LV. When these correlated constructs were estimated simultaneously, the HM coefficient reflected the unique component of enjoyment not shared with habit, commitment, or value. In the present model, that unique remainder was negative and non-significant. Substantively, this suggests that enjoyment may support intention mainly through its contribution to routine formation, persistence, and perceived worth of the learning experience, rather than through an independent direct pathway of its own. Once those more proximal engagement constructs were controlled, "pure enjoyment" did not add unique explanatory value and may have inverted slightly because of shared variance and suppression (Tamilmani et al., 2019; Deng & Yu, 2023).

This interpretation should nevertheless be treated cautiously. The current evidence supports a plausible suppression explanation, but stronger confirmation would require the additional sensitivity analyses requested by the reviewers. For the present study, the most defensible conclusion is that HM was positively related to BI at the bivariate level but did not retain a stable independent role once the broader engagement pathway was modeled.

5.4. Practical implications for Moodle-supported PML in the Maldives context

The findings suggest that increasing BI in this setting is best approached by designing for habit formation, strengthening commitment, and making learning value salient and experienced, not merely asserted (Ortiz de Guinea & Markus, 2009; Ain et al., 2016; San-Martín et al., 2020; Zacharis & Nikolopoulou, 2022). The practical implication is not that institutions should abandon support, expectancy-building, or social encouragement, but that these should be treated as enabling supports rather than as the primary levers of adoption in similar contexts.

Table 6 provides a prioritized action-oriented implementation map. These indicators are suggested as monitoring tools rather than as validated outcome measures and should therefore be piloted and adapted in context.

Table 6. Policy and Practice Priority Actions Derived from Results

Priority action	Stakeholder	Evidence strength	Cost/Effort	Timeline	Rationale linked to findings	Illustrative KPI(s)
Design a weekly “PML routine” workflow inside Moodle (fixed schedule, predictable task pattern, low-friction access)	Instructors + LMS admins	Strong	Medium	Short-term (4–8 weeks)	Builds habit (HT), the strongest unique predictor of BI	Weekly active user rate; proportion of students completing scheduled weekly PML tasks for 4–8 consecutive weeks
Embed quick feedback cycles (auto-graded quizzes + immediate explanations)	Instructors + LMS admins	Strong	Medium	Short-term	Increases experienced learning value (LV) and sustained engagement	Average quiz completion rate; average number of repeat practice attempts per student per week
Commitment supports: progress dashboards, goal-setting prompts, milestone badges tied to learning goals (not “fun only”)	LMS developers + instructors	Strong	Medium	Short–mid-term	Reinforces student commitment (SC); helps engagement persist until habitual	Goal completion rate; proportion of students maintaining engagement across a defined monitoring period (e.g., 6 weeks)
Align PML tasks to exam-relevant competencies and show “why this matters” explicitly	Instructors + curriculum leads	Strong	Low–Medium	Short-term	Strengthens learning value (LV); makes utility concrete	Student-reported perceived usefulness of task alignment; completion of competency-linked activities
Orientation and “first 2 weeks” scaffolding (walkthroughs + guided first tasks)	Administrators + instructors	Moderate	Low	Short-term	Helps move students from initial use to routine use (habit)	Orientation completion rate; first-two-week platform login consistency

					formation)	
Social influence interventions (peer study groups, mentor prompts)	Administrators	Weak–Moderate	Low	Mid-term	SI had a weaker unique role; used as support rather than a core lever	Participation in peer-support activities; proportion of students responding to mentor prompts
Infrastructure-only initiatives (devices/connectivity) as a standalone adoption strategy	Policymakers/administrators	Weak (as a standalone)	High	Long-term	FC correlates with BI but does not uniquely predict BI once engagement constructs are accounted for	Device/connectivity access coverage; service reliability indicators

Note. “Evidence strength” reflects alignment with significant predictors in the multivariate model (Table 4) and consistent bivariate support (Table 3). The KPI examples are illustrative and intended for short-cycle monitoring in pilot implementation rather than as definitive evaluative standards.

5.5. Contextual Interpretation and Transferability

The Maldives presents a distinctive educational and infrastructural context. Practical adoption may be affected by variability in connectivity reliability, students’ prior LMS exposure, and the extent to which digital learning is institutionalized across schools and pre-university programs (Azlifa & Saeed, 2021; Mumthaz, 2021). While facilitating conditions and expectancies matter in principle, the present results suggest that once baseline access is present, the differentiating drivers of intention are engagement-based: students intend to adopt Moodle-supported PML when it becomes routine, valued, and supported by commitment mechanisms (Ortiz de Guinea & Markus, 2009; Ain et al., 2016; San-Martín et al., 2020).

These implications should not be generalized without caution. The present findings come from a single institutional cohort with a modest response rate, and they therefore support recommendations for piloting and monitored implementation in similar contexts rather than universal prescriptions across all institutions. Settings with substantially lower access or weaker institutional embedding of Moodle may see FC and EE play larger roles, consistent with barrier-and-enabler patterns reported in other constrained contexts (Al-Azawei et al., 2016; Akwene, 2024). Conversely, contexts with mature LMS integration may show the same pattern observed here: expectancy constructs become necessary but not sufficient, while habit- and value-based engagement dominate (Zacharis & Nikolopoulou, 2022; Moorthy et al., 2019).

The low mean for SI also has context-specific policy implications. If social influence is comparatively weak in this sample, institutional strategies in similar settings may need to rely less on one-way encouragement campaigns and more on embedding Moodle-supported PML into routine instructional design, feedback cycles, and visible academic relevance. In other words, students may need to

experience the platform as educationally useful and habit-supporting rather than simply being told that they should use it.

5.6. Methodological Considerations: OLS Versus SEM and Statistical Inference

The study used OLS regression to estimate the association between the predictor set and BI because it provides transparent coefficient interpretation and aligns with the study's objective of identifying key determinants of intention. Regression diagnostics indicated acceptable collinearity and residual behaviour for inference (Table 4). However, UTAUT-type models are commonly estimated with SEM, which explicitly models latent variables and measurement error (Venkatesh et al., 2012; Venkatesh et al., 2016). SEM may therefore yield more precise parameter estimates and allow simultaneous testing of the measurement model and structural relationships (Venkatesh et al., 2012).

The choice of OLS in the present study should therefore be understood as a pragmatic analytical decision rather than as a claim that OLS is the optimal method for all UTAUT-based research. It was selected because it supports direct interpretation of coefficients for applied audiences and is compatible with the study's main objective of identifying relative predictor salience. To address the issue of multiple testing, the regression model is treated as the primary multivariate test, and bivariate correlations as descriptive context rather than as independent hypothesis tests.

5.7. Limitations and Future Research

Several limitations should guide interpretation. First, the cross-sectional design limits causal inference; behavioural intention is a proximal predictor and does not confirm actual Moodle usage or learning performance outcomes (Venkatesh et al., 2012). Second, reliance on self-report measures introduces the potential for common-method bias. Third, the achieved sample was based on voluntary participation and a 21.5% response rate, which limits generalizability beyond the sampled cohort and raises the possibility of non-response bias. Fourth, while collinearity diagnostics were acceptable, moderate intercorrelations among motivational constructs likely contributed to suppression patterns (e.g., HM) (Dwivedi et al., 2019; Tamilmani et al., 2019). Fifth, if the measurement model continues to rely primarily on EFA rather than a full CFA-based validation strategy, conclusions about construct distinctiveness should remain appropriately cautious.

Future research should (i) measure actual Moodle usage logs and learning outcomes, (ii) adopt longitudinal designs to track habit formation over time, (iii) test potential moderators such as prior Moodle experience and connectivity constraints, and (iv) estimate the model using SEM for latent-variable inference where sample size permits (Venkatesh et al., 2012; Venkatesh et al., 2016). Future studies should also compare the explanatory performance of extended UTAUT2 models across different educational contexts to clarify when value- and commitment-based constructs outperform expectancy-based constructs as predictors of sustained adoption.

6. Conclusion

This study set out to understand what most strongly aligns with pre-university students' intention to adopt Moodle-supported Personalized Mathematics Learning (PML) in the Maldives. The evidence points to a clear pattern: intention is not driven primarily by whether students think the platform is useful or easy in the abstract, but by whether the learning experience becomes meaningful, sustained, and routine. When students perceive real learning value in PML, feel personally committed to staying engaged, and develop habitual patterns of using Moodle for their mathematics work, intention is markedly stronger. At the same time, the findings also suggest that the canonical UTAUT2 predictors, performance expectancy, effort expectancy, social influence, and facilitating condition, may function more as supportive background conditions than as the strongest unique drivers of intention in this

learning-specific context. This implies that educational technology adoption models may need to place greater emphasis on value, persistence, and routinised engagement when the focus is sustained learning use rather than initial acceptance alone.

PML supported by Moodle should be perceived as more than the introduction of a certain platform and the expectation that students will use it. The crucial practical consideration is whether the learning experience is structured so that students can revisit it periodically, with clear tasks, visible progress, and consistent feedback that supports persistence. Once these conditions have been established, it becomes easier to maintain further use and may slowly evolve into habitual interaction. For institutions, this means that adoption is unlikely to be secured through a single initial orientation or one-time sensitisation effort; rather, it depends on the consistent integration of Moodle-supported PML into teaching, practice routines, and feedback processes that sustain continued student use over time.

These findings should also be interpreted with caution. The study is cross-sectional and focuses on behavioural intention rather than actual usage or learning outcomes. The results, however, do not imply causal influences but instead suggest associations. The evidence can be further enhanced in the future by monitoring students' use of Moodle over time, including learning analytics and log data, and by examining whether maintaining Moodle-supported PML use is related to improvements in mathematics learning outcomes.

These limitations notwithstanding, the study has a definite implication for LMS-based personalised learning in similar contexts: the intention to adopt is most pronounced when students find the value of learning evident, are ready to commit to it, and establish a repeated usage pattern that becomes a habit. In this respect, PML supported by Moodle is effective not just because of its presence, but also because of the experience gained through students' normal learning activities. For MNU and similar Maldives higher education contexts, the most practical next step is to pilot implementation strategies that strengthen routine use, visible learning value, and sustained commitment, while monitoring student engagement over short and medium time frames to determine which actions most effectively support continued adoption.

Acknowledgment

The first author sincerely thanks the Islamic Development Bank (IsDB) for awarding her the Merit PhD Scholarship (Grant No. 2021-580235).

Conflict of Interest

All authors declare no competing interests or potential conflicts of interest that could compromise the integrity of the research.

References

- Abbad, M. M. M. (2021). Using the UTAUT model to understand students' usage of e-learning systems in developing countries. *Education and Information Technologies*, 26(6), 7205–7224. <https://doi.org/10.1007/s10639-021-10573-5>
- Adam, A. S. (2017). A Framework for Seeking the Connections between Technology, Pedagogy, and Culture: A study in the Maldives. *Journal of Open, Flexible and Distance Learning*, 21(1), 35–51. <https://doi.org/10.61468/jofdl.v21i1.280>
- Ain, N. U., Kaur, K., & Waheed, M. (2016). The influence of learning value on learning management system use: An extension of UTAUT2. *Information Development*, 32(5), 1306–1321. <https://doi.org/10.1177/0266666915597546>
- Akwene, G. C. (2024). Adoption of Learning Management Systems in Higher Education: A Qualitative Analysis of Enabling Factors. *International Journal of Scientific Research and Management (IJSRM)*, 12(02), 3229–3237. <https://doi.org/10.18535/ijrm/v12i02.el09>
- Al-Azawei, A., Parslow, P., & Lundqvist, K. (2016). Barriers and opportunities of e-learning implementation in Iraq: A case of public universities. *International Review of Research in Open and Distributed Learning*, 17(5), 126–146. <https://doi.org/10.19173/irrodl.v17i5.2501>
- Alamri, H., Lowell, V., Watson, W., & Watson, S. L. (2020). Using personalized learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation. *Journal of Research on Technology in Education*, 52(3), 322–352. <https://doi.org/10.1080/15391523.2020.1728449>
- Azizi, S. M., Roozbahani, N., & Khatony, A. (2020). Factors affecting the acceptance of blended learning in medical education: application of UTAUT2 model. *BMC Medical Education*, 20(1), 1–9. <https://doi.org/10.1186/s12909-020-02302-2>
- Azlifa, M., & Saeed, F. (2021). The challenges to digitalization of schools in the Maldives 58 The Challenges to Digitalization of Schools in the Maldives. *International Journal of Social Research and Innovation* |, 5(2).
- Bervell, B., & Umar, I. N. (2017). Validation of the UTAUT model: Re-considering non-linear relationships of exogenous variables in higher education technology acceptance research. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(10), 6471–6490. <https://doi.org/10.12973/ejmste/78076>
- Bradley, V. M. (2020). Learning Management System (LMS) Use with Online Instruction. *International Journal of Technology in Education*, 4(1), 68. <https://doi.org/10.46328/ijte.36>
- Brown, S., Murphy, L., & Hammond, K. (2022). Learning Management System Adoption by Academics. *Journal of Open, Flexible and Distance Learning*, 25(2), 55–65. <https://doi.org/10.61468/jofdl.v25i2.477>
- Chávez Herting, D., Cladellas Pros, R., & Castelló Tarrida, A. (2023). Habit and social influence as determinants of PowerPoint use in higher education: A study from a technology acceptance approach. *Interactive Learning Environments*, 31(1), 497–513. <https://doi.org/10.1080/10494820.2020.1799021>
- Chen, M., Wang, X., Wang, J., Zuo, C., Tian, J., & Cui, Y. (2021). Factors Affecting College Students' Continuous Intention to Use Online Course Platform. *SN Computer Science*, 2(2), 1–11.

<https://doi.org/10.1007/s42979-021-00498-8>

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)

Deng, X., & Yu, Z. (2023). An extended hedonic motivation adoption model of TikTok in higher education. *Education and Information Technologies*, 28(10), 13595–13617. <https://doi.org/10.1007/s10639-023-11749-x>

Dindar, M., Suorsa, A., Hermes, J., Karppinen, P., & Näykki, P. (2021). Comparing technology acceptance of K-12 teachers with and without prior experience of learning management systems: A Covid-19 pandemic study. *Journal of Computer Assisted Learning*, 37(6), 1553–1565. <https://doi.org/10.1111/jcal.12552>

Dokhanian, S., Roustapisheh, N., Heidari, S., & Rezvani, S. (2022). The Effectiveness of System Quality, Habit, and Effort Expectation on Library Application Use Intention: The Mediating Role of Perceived Usefulness, Perceived Ease of Use, and User Satisfaction. *International Journal of Business Information Systems*, 1(1), 1. <https://doi.org/10.1504/ijbis.2022.10049515>

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model. *Information Systems Frontiers*, 21(3), 719–734. <https://doi.org/10.1007/s10796-017-9774-y>

El-Masri, M., & Tarhini, A. (2017). Factors affecting the adoption of e-learning systems in Qatar and USA: Extending the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). *Educational Technology Research and Development*, 65(3), 743–763. <https://doi.org/10.1007/s11423-016-9508-8>

Fatmi, N., Muhammad, I., Muliana, M & Nasrah, S. (2021). The Utilization of Moodle-Based Learning Management System (LMS) in Learning Mathematics and Physics to Students' Cognitive Learning Outcomes. *International Journal for Educational and Vocational Studies*, 3(2), 155. <https://doi.org/10.29103/ijevs.v3i2.4665>

Feitosa de Moura, V., Alexandre de Souza, C., & Noronha Viana, A. B. (2021). The use of Massive Open Online Courses (MOOCs) in blended learning courses and the functional value perceived by students. *Computers and Education*, 161(November 2020). <https://doi.org/10.1016/j.compedu.2020.104077>

Gamage, S. H. P. W., Ayres, J. R., & Behrend, M. B. (2022). A systematic review on trends in using Moodle for teaching and learning. *International Journal of STEM Education*, 9(1). <https://doi.org/10.1186/s40594-021-00323-x>

Gansser, O. A., & Reich, C. S. (2021). A new acceptance model for artificial intelligence with extensions to UTAUT2: An empirical study in three segments of application. *Technology in Society*, 65, 101535. <https://doi.org/10.1016/j.techsoc.2021.101535>

Gardašević, D., Banjević, K., Nastasić, A., & Rošulj, D. (2023). Student Achievements and Their Attitudes Toward Course Design - a Correlation Analysis. *Journal of Process Management and New Technologies*, 11(1–2), 34–56. <https://doi.org/10.5937/jpmnt11-44534>

Guine, A. O. de & M. L. M. (2009). Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology Use Author(s): Ana Ortiz de Guinea and. *Source: MIS Quarterly*, 33(3), 433–444.

- Hewavitharana, T., Nanayakkara, S., Perera, A., & Perera, P. (2021). Modifying the unified theory of acceptance and use of technology (UTAUT) model for the digital transformation of the construction industry from the user perspective. *Informatics*, 8(4). <https://doi.org/10.3390/informatics8040081>
- Hooda, A., Gupta, P., Jeyaraj, A., Giannakis, M., & Dwivedi, Y. K. (2022). The effects of trust on behavioral intention and use behavior within e-government contexts. *International Journal of Information Management*, 67(August), 102553. <https://doi.org/10.1016/j.ijinfomgt.2022.102553>
- Huang, F., Teo, T., & Scherer, R. (2022). Investigating the antecedents of university students' perceived ease of using the Internet for learning. *Interactive Learning Environments*, 30(6), 1060–1076. <https://doi.org/10.1080/10494820.2019.1710540>
- Hüseyin, U., Murat, T., & Ezgi, P. Y. (2020). The effects of the authentic learning approach with a course management system (moodle) on students mathematics success and online authentic learning self-efficacy. *Educational Research and Reviews*, 15(11), 679–689. <https://doi.org/10.5897/err2020.4087>
- Ismail, S. N., Hamid, S., Ahmad, M., Alaboudi, A., & Jhanjhi, N. (2021). Exploring students engagement towards the learning management system (LMS) using learning analytics. *Computer Systems Science and Engineering*, 37(1), 73–87. <https://doi.org/10.32604/CSSE.2021.015261>
- Kapsargina, S., & Olentsova, J. (2020). Experience of Using LMS Moodle in the Organization of Independent Work of Bachelors in Teaching a Foreign Language. *Proceedings of the International Scientific Conference "Far East Con" (ISCFEC 2020)*, 128(Iscfec), 537–544. <https://doi.org/10.2991/aebmr.k.200312.076>
- Khan, W. A., & Abideen, Z. U. (2023). Effects of behavioural intention on usage behaviour of digital wallet: the mediating role of perceived risk and moderating role of perceived service quality and perceived trust. *Future Business Journal*, 9(1). <https://doi.org/10.1186/s43093-023-00242-z>
- Khechine, H., Raymond, B., & Augier, M. (2020). The adoption of a social learning system: Intrinsic value in the UTAUT model. *British Journal of Educational Technology*, 51(6), 2306–2325. <https://doi.org/10.1111/bjet.12905>
- Kim, J., & Lee, K. S. S. (2022). Conceptual model to predict Filipino teachers' adoption of ICT-based instruction in class: using the UTAUT model. *Asia Pacific Journal of Education*, 42(4), 699–713. <https://doi.org/10.1080/02188791.2020.1776213>
- Kumar, J. A., & Bervell, B. (2019). Google Classroom for mobile learning in higher education: Modelling the initial perceptions of students. *Education and Information Technologies*, 24(2), 1793–1817. <https://doi.org/10.1007/s10639-018-09858-z>
- Lai, Y., Saab, N., & Admiraal, W. (2022). University students' use of mobile technology in self-directed language learning: Using the integrative model of behavior prediction. *Computers and Education*, 179(November 2021), 104413. <https://doi.org/10.1016/j.compedu.2021.104413>
- Lisnani, L., Putri, R. I. I., Zulkardi, & Somakim. (2020). Designing Moodle features as e-learning for learning mathematics in COVID-19 pandemic. *Journal of Physics: Conference Series*, 1657(1). <https://doi.org/10.1088/1742-6596/1657/1/012024>
- Liu, H., Lin, C. H., Zhang, D., & Zheng, B. (2018). Chinese Language Teachers' Perceptions of Technology and Instructional Use of Technology: A Path Analysis. *Journal of Educational Computing Research*, 56(3), 396–414. <https://doi.org/10.1177/0735633117708313>
- Maghrour Zefreh, M., Edries, B., & Esztergár-Kiss, D. (2023). Understanding the antecedents of

hedonic motivation in autonomous vehicle technology acceptance domain: A cross-country analysis. *Technology in Society*, 74(July). <https://doi.org/10.1016/j.techsoc.2023.102314>

Major, L., Francis, G. A., & Tsapali, M. (2021). The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. *British Journal of Educational Technology*, 52(5), 1935–1964. <https://doi.org/10.1111/bjet.13116>

Matosas-López, L., Aguado-Franco, J. C., & Gómez-Galán, J. (2019). Constructing an instrument with behavioral scales to assess teaching quality in blended learning modalities. *Journal of New Approaches in Educational Research*, 8(2), 142–165. <https://doi.org/10.7821/naer.2019.7.410>

Mlotshwa, N., Tunjera, N., & Chigona, A. (2020). Integration of moodle into the classroom for better conceptual understanding of functions in mathematics. *South African Journal of Education*, 40(3), 1–14. <https://doi.org/10.15700/saje.v40n3a1570>

Mohamed, N. (2020). Relationship between formative assessments given on moodle and self-regulatory learning skills among blended learning students at the maldives national university. *ACM International Conference Proceeding Series, Part F16898*, 48–54. <https://doi.org/10.1145/3446590.3446598>

Molins, L. L., & García, E. C. (2023). How to Embed SRL in Online Learning Settings? Design Through Learning Analytics and Personalized Learning Design in Moodle. *Journal of New Approaches in Educational Research*, 12(1), 120–138. <https://doi.org/10.7821/naer.2023.1.1127>

Moorthy, K., Yee, T. T., T'ing, L. C., & Kumaran, V. V. (2019). Habit and hedonic motivation are the strongest influences in mobile learning behaviours among higher education students in Malaysia. *Australasian Journal of Educational Technology*, 35(4), 174–191. <https://doi.org/10.14742/ajet.4432>

Muhfahroyin, M., & Sujarwanta, A. (2023). Student's commitments and preferences in online learning. *Journal of Education and Learning*, 17(2), 326–334. <https://doi.org/10.11591/edulearn.v17i2.20698>

Müller, C., & Mildenerger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34(November 2020), 100394. <https://doi.org/10.1016/j.edurev.2021.100394>

Mumthaz, F. (2021). Psychological Readiness for Mobile Learning Among Maldivian Students at Higher Education Institutions. *GATR Global Journal of Business Social Sciences Review*, 9(3), 224–232. [https://doi.org/10.35609/gjbssr.2021.9.3\(4\)](https://doi.org/10.35609/gjbssr.2021.9.3(4))

Mursyidah, H., Hermoyo, R. P., & Suwaibah, D. (2021). Does flipped learning method via MOODLE can improve outcomes and motivation of discrete mathematics learning during COVID-19 pandemic?. In *Journal of Physics: Conference Series* (Vol. 1720, No. 1, p. 012007). IOP Publishing.

Nikolopoulou, K., Gialamas, V., & Lavidas, K. (2021). Habit, hedonic motivation, performance expectancy and technological pedagogical knowledge affect teachers' intention to use mobile internet. *Computers and Education Open*, 2(March), 100041. <https://doi.org/10.1016/j.caeo.2021.100041>

Oudeyer, P. Y., & Kaplan, F. (2009). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 3(NOV), 1–14. <https://doi.org/10.3389/neuro.12.006.2007>

Özsungur, F. (2019). The effects of technology acceptance and use behaviour on women's entrepreneurship motivation factors. *Asia Pacific Journal of Innovation and Entrepreneurship*, 13(3), 367–380. <https://doi.org/10.1108/apjie-09-2019-0070>

Peneda, F. P., Placigo, M. R., Esporas, R. R., & Ticoy, T. T. (2023). Insights on the Implementation of Learning Management System in Mathematics Class. *International Journal of Research Publications*, 117(1). <https://doi.org/10.47119/ijrp1001171120234447>

Perez-Suay, A., Van Vaerenbergh, S., Diago, P. D., Pascual-Venteo, A. B., & Ferri, F. J. (2023). Data-Driven Modeling Through the Moodle Learning Management System: An Empirical Study Based on a Mathematics Teaching Subject. *Revista Iberoamericana de Tecnologías Del Aprendizaje*, 18(1), 19–27. <https://doi.org/10.1109/RITA.2023.3250434>

Pramod, D., Vijayakumar Bharathi, S., & Patil, K. (2024). Generative Artificial Intelligence for Personalized Elderly Care through the lens of the Hedonic Motivation System Adoption Model. *Proceedings - 2nd International Conference on Advancement in Computation and Computer Technologies, InCACCT 2024*, 105–110. <https://doi.org/10.1109/InCACCT61598.2024.10551017>

Pratama, R., Aisyah, S. A., Putra, A. M., Sirodj, R. A., & Afgan, M. W. (2023). Correlational Research. *JiIP - Jurnal Ilmiah Ilmu Pendidikan*, 6(3), 1754–1759. <https://doi.org/10.54371/jiip.v6i3.1420>

Raza, S. A., Qazi, W., Khan, K. A., & Salam, J. (2021). Social Isolation and Acceptance of the Learning Management System (LMS) in the time of COVID-19 Pandemic: An Expansion of the UTAUT Model. *Journal of Educational Computing Research*, 59(2), 183–208. <https://doi.org/10.1177/0735633120960421>

Salloum, S. A., & Shaalan, K. (2019). Factors Affecting Students' Acceptance of E-Learning System in Higher Education Using UTAUT and Structural Equation Modeling Approaches. In *Advances in Intelligent Systems and Computing* (Vol. 845). Springer International Publishing. https://doi.org/10.1007/978-3-319-99010-1_43

San-Martín, S., Jiménez, N., Rodríguez-Torrico, P., & Piñeiro-Ibarra, I. (2020). The determinants of teachers' continuance commitment to e-learning in higher education. *Education and Information Technologies*, 25(4), 3205–3225. <https://doi.org/10.1007/s10639-020-10117-3>

Shemshack, A., & Spector, J. M. (2020). A systematic literature review of personalized learning terms. *Smart Learning Environments*, 7(1). <https://doi.org/10.1186/s40561-020-00140-9>

Sher, V., Hatala, M., & Gašević, D. (2020). Analyzing the consistency in within-activity learning patterns in blended learning. *ACM International Conference Proceeding Series*, 1–10. <https://doi.org/10.1145/3375462.3375470>

Tamilmani, K., Rana, N. P., Prakasam, N., & Dwivedi, Y. K. (2019). The battle of Brain vs. Heart: A literature review and meta-analysis of “hedonic motivation” use in UTAUT2. *International Journal of Information Management*, 46(February), 222–235. <https://doi.org/10.1016/j.ijinfomgt.2019.01.008>

Turnbull, D., Chugh, R., & Luck, J. (2021). Transitioning to E-Learning during the COVID-19 pandemic: How have Higher Education Institutions responded to the challenge? *Education and Information Technologies*, 26(5), 6401–6419. <https://doi.org/10.1007/s10639-021-10633-w>

Venkatesh, V., Morris, M. G., Hall, M., Davis, G. B., Davis, F. D., & Walton, S. M. (2003). *User Acceptance of Information Technology: Toward a Unified View*. 27(3), 425–478.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly: Management Information Systems*, 36(1), 157–178. <https://doi.org/10.2307/41410412>

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). Unified theory of acceptance and use of technology: A

synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376. <https://doi.org/10.17705/1jais.00428>

Waterschoot, J., Vansteenkiste, M., & Soenens, B. (2019). The effects of experimentally induced choice on elementary school children's intrinsic motivation: The moderating role of indecisiveness and teacher–student relatedness. *Journal of Experimental Child Psychology*, 188. <https://doi.org/10.1016/j.jecp.2019.104692>

Wilkins, S., Butt, M. M., Kratochvil, D., & Balakrishnan, M. S. (2016). The effects of social identification and organizational identification on student commitment, achievement and satisfaction in higher education. *Studies in Higher Education*, 41(12), 2232–2252. <https://doi.org/10.1080/03075079.2015.1034258>

Xie, H., Kitcharoen, K., Leelakasemsant, C., Mechankara Varghese, M., & Au-Gsb E, /. (2022). Adoption of online classes during COVID-19: An institution's investigation on perception & behavioral intention. *Assumptionjournal.Au.Edu*, 15(2), 159–168.

Yu, C. W., Chao, C. M., Chang, C. F., Chen, R. J., Chen, P. C., & Liu, Y. X. (2021). Exploring Behavioral Intention to Use a Mobile Health Education Website: An Extension of the UTAUT 2 Model. *SAGE Open*, 11(4). <https://doi.org/10.1177/21582440211055721>

Zacharis, G., & Nikolopoulou, K. (2022). Factors predicting University students' behavioral intention to use eLearning platforms in the post-pandemic normal: an UTAUT2 approach with 'Learning Value.' *Education and Information Technologies*, 27(9), 12065–12082. <https://doi.org/10.1007/s10639-022-11116-2>

Zheng, H., Han, F., Huang, Y., Wu, Y., & Wu, X. (2025). Factors influencing behavioral intention to use e-learning in higher education during the COVID-19 pandemic: A meta-analytic review based on the UTAUT2 model. *Education and Information Technologies*, 30, 12015–12053. <https://doi.org/10.1007/s10639-024-13299-2>

Organizational Identification of Teachers in Private Higher Education: Insights from a Mixed-Methods Study

ABSTRACT

Authors: Xiuqiang Feng¹, Rahimah Jamaluddin^{1*} and Mohd Hazwan Mohd Puad¹

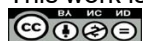
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:

imah_upm@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Teachers' organizational identification is particularly important for maintaining institutional stability and development. However, due to the limitation in governance structure, resource availability, and employment models, private higher education institutions confront higher challenges than those of public institutions. To strengthen teachers' organizational identification, this study aims to propose effective strategies by examining its current status, group differences, and relevant contextual influences. Based on a three-factor framework encompassing existence, belongingness, and success, the ERG theory is adopted and a questionnaire is developed to collect quantitative data involving 268 teachers. Also, semi-structured qualitative interviews of 11 teachers are complemented to explore an in-depth understanding. Results show that teachers' organizational identification in private higher education institutions is generally at a moderate level, and the identification degree in descending order is belongingness, success, and existence. Note that age, educational level, academic title, and establishment status may bring minor differences, whereas gender, full-time status, or administrative duties have no significant impact. The contextualize and pattern explanations can be drawn from interviews, as well as additional insights for the development of targeted improvement strategies. Lastly, this study proposes three key management steps: implementing coordinated strategies to enhance existence identification, building a supportive climate for belongingness identification, and expanding resources to boost success identification.

Keywords: Organizational Identification, Private Higher Education, Teachers, ERG Theory, Mixed Methods.

1. Introduction

Higher education is gradually moving towards a trend of diversified funding sources and market oriented management models. Unlike public institutions that have stable public financial support and well established governance structures, private colleges and universities are more directly exposed to market competition, resource fluctuations, and employment uncertainties (Buckner, 2025). They rely more heavily on effective institutional management and organizational support to survive and achieve sustainable development. As one of the core competitive advantages of higher education institutions, teachers' stability and commitment have become essential to achieving broader strategic goals. In the context of limited resources in private higher education institutions, enhancing teachers' identification with the organization has become an important governance strategy (Weisman et al., 2023; Christopher & Divya, 2024).

Organizational identification is generally regarded as the cognitive and emotional psychological bond between individuals and their organization, which develops a sense of belonging and incorporates it into their own self-perception (Ashforth & Mael, 1989; Christopher & Divya, 2024; Weisman et al., 2023). For private higher education institutions that have limited resources and are under the pressure of market competition, improving teachers' organizational identification is not only a psychological requirement, but also a crucial management task. If institutional decision-makers lack a thorough understanding of these mechanisms, their intervention strategies may fail to meet teachers' needs, which could lead to organizational disintegration (Hu & Ghasemy, 2024). However, despite its considerable strategic importance, empirical evidence remains limited on how to manage organizational identification among teachers in private colleges and universities. This gap may expose institutions' development to substantial risks. Therefore, it is important to investigate teachers' organizational identification for improving management practices and ensuring the long-term stability and development of private higher education institutions.

In order to provide evidence-based foundations for institutional leaders, this study focuses on teachers in China's private higher education institutions and explores their organizational identification from a context-sensitive perspective. As one of the largest and most rapidly evolving private higher education systems globally, China serves as a representative case. In particular, Chinese private higher education is undergoing a transformation from rapid scale expansion to quality development, teachers' organizational identification becomes more critical to institutional success (State Council of the PRC, 2021; Hu & Ghasemy, 2024; Peng, 2025).

This study also seeks to provide valuable insights for other higher education systems beyond China. Similar patterns can be observed in countries such as Malaysia, Indonesia, Vietnam, and the Philippines, where private institutions are playing an increasingly important role in expanding the accessibility of higher education. In these contexts, private higher education institutions often face similar challenges related to resource constraints, governance capabilities, and teacher retention (Nair et al., 2024; Mohamad & Majid, 2025; Wiyono et al., 2025). Therefore, the findings of this study may provide useful references for institutional leaders seeking to enhance organizational support and improve institutional stability.

To guide the research, this study explores the following research questions (RQ):

RQ1: What is the current status and the detailed dimensional structure of teachers' organizational identification in private higher education?

RQ2: How does teachers' organizational identification differ across demographic and positional characteristics?

RQ3: What contextual factors help explaining variations in teachers' organizational identification in private higher education?

RQ4: What management and leadership strategies can be derived to enhance teachers' organizational identification and support institutional development?

2. Literature Review

2.1. Importance of Organizational Identification in Private Higher Education

Among teachers in higher education, organizational identification has been linked to a wide set of outcomes. According to prior studies, it correlates positively with organizational citizenship behavior, work engagement, and performance (Sidorenkov & Muchinsky, 2023; Khanal & Arora, 2024; Pugliese et al., 2024). Stronger identification is associated with greater discretionary effort in teaching reform, curriculum development, talent cultivation, and teamwork (Li, 2024; Pugliese et al., 2024). It also shows a positive association with job satisfaction, emotional investment, and career commitment, which can relieve stress, enhance well-being and resilience (Jun et al., 2023; Wang, 2025). Furthermore, it has a negative correlation with turnover intention, indicating that it plays a more significant role in promoting long-term retention and stability (Feng et al., 2025; Yang et al., 2025). Overall, prior studies have reported largely consistent positive outcomes, but relatively limited attention has been given to how organizational identification develops among teachers in higher education.

In recent years, research on organizational identification in higher education has been continuously expanding in various contexts. For instance, Chen et al. (2025) analysed survey data from 283 higher education faculty members and examined organizational identification across three dimensions: cohesion, group influence, and group relationships. Similarly, Borlaug et al. (2023) investigated organizational identification among academic staff during the merger of higher education institutions using a mixed-method design, showing that institutional restructuring and access to organizational resources played an important role in shaping staff identification.

Although these efforts are constantly increasing, research specifically targeting teachers at private colleges and universities remains limited. Firstly, the reliance on a undifferentiated sample of higher education obscures the unique governance structure specific to private institutions, preventing leaders from obtaining governance solutions tailored to this particular field. Secondly, the scales or questionnaires originally designed for corporate or general organizational contexts cannot accurately reflect the unique concepts of private higher education, resulting in a disconnection between the measurement results and the actual situation, and thereby misleading policy-making. Thirdly, the fragmentary understanding of demographics hinders the formulation of targeted management strategies for different groups of teachers. Finally, the excessive reliance on descriptive indicators fails to examine the actual experiences of teachers, thereby masking the underlying reasons that lead to teachers' feelings of alienation.

These research gaps have brought continuous challenges to institutional management, as the lack of evidence tailored to specific situations may lead to administrative decisions not fully aligning with the professional needs of teachers.

2.2. ERG Theory

In 1969, based on Maslow's need hierarchy theory from the perspective of humanism, Clayton P. Alderfer proposed a new needs theory – ERG theory (Alderfer, 1969; 1972). Alderfer (1972) proposed three core categories of human needs, including existence (E), relatedness (R), and growth (G). The existence needs refer to the physiological and material requirements, including salary, benefits, working conditions, safety and stability. The relatedness needs, on the other hand, are generated through interaction in life and work, such as interpersonal relationships, teamwork, fairness in procedures and distribution, respect, and a sense of belonging within the organization. The driving force for growth needs stems from the desire for self-improvement and development. This desire manifests as an inherent spirit of enterprise, which is a strong urge to progress, to realize personal value, and to achieve success (Yang & Ling, 2023).

The ERG theory is based on four assumptions. First, satisfaction of lower-level needs is not a prerequisite for the pursuit of higher-level needs. Individuals can simultaneously pursue multiple levels, rather than following a strictly linear sequence. For instance, a teacher whose compensation is not yet fully satisfactory may nonetheless strongly seek academic development and professional accomplishments. Second, the frustration–regression mechanism: when the pursuit of higher-level needs (e.g., growth) is blocked, individuals may turn to lower-level needs (e.g., existence or relatedness) for compensatory satisfaction to maintain motivational equilibrium. Third, needs are not fixed. The dominant needs at a given stage can shape behavior for a period. Once certain needs are met, the person may proceed to higher-level needs or remain at the current stage due to contextual constraints. Fourth, even after partial satisfaction, the expectation level for a given need may continue to rise rather than diminish (Alderfer, 1972; Yang & Ling, 2023).

Within private colleges and universities, teachers likewise require multifaceted need satisfaction from the organization to activate development motives and form stable organizational attitudes. In ERG terms, the three need categories can operate concurrently, jointly shaping work experiences and identity linkage. When existence needs (e.g., compensation and job stability) are adequately met, attention is more readily directed toward relatedness (e.g., collegial collaboration) and growth (e.g., development opportunities). Conversely, when growth channels are obstructed or organizational support is lacking, frustration–regression can occur, with teachers seeking compensation along existence and relatedness dimensions (e.g., stronger concern for pay equity, and team climate). ERG theory therefore does not treat needs as rigid. It highlights dynamic reversibility and stage sensitivity, features that fit the fluctuating need profiles of teachers in private colleges and universities across career stages, positions, and institutional arrangements.

In this study, ERG theory is adopted as the theoretical basis for explaining how need-related feelings shape teachers' organizational identification in private higher education. However, organizational identification has been predominantly studied through the lens of Social Identity Theory (Tajfel & Turner, 1979), Self-categorization Theory (Turner et al., 1987), or Person-organization Fit Theory (Kristof, 1996). These perspectives emphasize self-categorization and psychological attachment to the organization. While they explain the social and cognitive processes underlying identification, ERG theory focuses on how the satisfaction of existence, relatedness, and growth needs influences employees' attitudes toward their organization. In the context of private colleges and universities, teachers' perceptions of compensation, interpersonal atmosphere, and development opportunities are closely related to their daily work experiences. Therefore, a needs-based framework provides a practical perspective for understanding the formation of organizational identification. By linking teachers' needs to existence, relatedness, and growth, institutions can better identify deficiencies in organizational support and develop targeted strategies related to workload, professional development, and participatory governance.

2.3. Three-Factor Model of Organizational Identification

On the basis of ERG theory, Wang (2005) proposed a three-factor model of organizational identification tailored to Chinese organizational settings, comprising existence identification, belongingness identification, and success identification. Based on the inherent logic formed by elements such as identity, institutions and resources in the organizational context of China, this model divides the members' sense of identification with the organization into three interrelated but distinct psychological perceptions. The existence identification refers to the subjective perception and recognition of members regarding the organizational support that can provide personal life and work security, including material guarantees, salary and benefits, and job stability. The belongingness identification refers to the emotional connection and sense of belonging that is formed in various situations such as interpersonal interactions, organizational culture, procedural fairness, leadership support, and teamwork. The success identification refers to the positive psychological experience and value cognition that individuals develop when they recognize improvements in their own abilities, access to resources, and career advancement. While primarily validated within Chinese organizational settings, this three-factor model

has gained widespread recognition in subsequent research and has shown good discriminant power and structural validity (Li et al., 2008; Yang et al., 2019; Li et al., 2023; Feng et al., 2025).

The reason why this model has a good theoretical correlation with the ERG theory lies in the fact that they share a high degree of consistency in terms of structural meaning and mechanism. The existence, relatedness, and growth in the ERG theory can correspond to the recognition of existence, belongingness, and success identification. Taking into account the professional ecology and organizational characteristics of teachers in private colleges and universities, the connotations of these three factors can be specifically expressed as follows. Existence identification emphasizes the sense of being properly secured by material and institutional support, such as the adequacy and fairness of pay and benefits, the stability of posts and contracts, and the reliability of working conditions and institutional safeguards, which corresponds to the existence needs in ERG theory. Belongingness identification emphasizes the relational and affective linkage formed in organizations, including fairness in procedures and processes, leadership support and respect, collegial trust and collaboration, and consistency between organizational culture and values, which corresponds to the relatedness needs in ERG theory. Success identification emphasizes the experience of achievement and development under organizational support, such as the provision of academic and teaching resources, training and promotion channels, external cooperation and platform opportunities, and recognition and incentive mechanisms, which corresponds to the growth needs in ERG theory.

Table 1. ERG Theory and the Three-Factor Model of Organizational Identification

ERG components	Three-factor model	Conceptual meaning
Existence	Existence identification	Perceived security and material support provided by the institution
Relatedness	Belongingness identification	Emotional connection and sense of belonging formed through organization interaction
Growth	Success identification	Perceived opportunities for professional development and achievement within the organization

Thus, the analytical framework of this study is established, as illustrated in Table 1. ERG theory serves as the underlying logic of needs satisfaction and motivational activation, and the three-factor model serves as the operational vehicle for the structured presentation of organizational identification. The two are integrated to explain the status and difference characteristics of organizational identification among teachers in private higher education, and to provide a theoretical basis for subsequent contextual interpretation and governance recommendations. Their integration enables to gain a deeper understanding of how motivational needs, leadership practices, and governance arrangements collectively shape the identification patterns of teachers in private higher education institutions.

3. Methodology

Based on a review of the relevant literature on teacher organizational identification, this study adopted a mixed-method design, combining questionnaire surveys with semi-structured interviews. The questionnaire survey comprehensively reflects the level and structure of teachers' organizational identification, as well as the group differences. The interviews are used to explore potential factors and propose context-specific suggestions. Through the mixed-method design, this study not only captured the breadth of the identification patterns through quantitative indicators, but also revealed the depth of personal experiences through qualitative narratives. Thus, the proposed strategies are both empirically grounded and context sensitive. Ethical approval for the study was obtained from the Ethics Review Committee of the Asia-Pacific Social Innovation Research Institute (Approval No. APSIRI-ERC-2025-1F78).

3.1. Questionnaire Survey

3.1.1. Instrument

The questionnaire on teacher organizational identification in private higher education is divided into two parts. The first part is used to measure organizational identification. This part draws on the Organizational Identification Scale developed by Wang (2005), which has been adapted based on previous research and initial teacher interviews, to suit the actual situation of private colleges and universities. The second part of the questionnaire collects basic information about teachers. Expert guidance is considered crucial for ensuring the applicability, completeness and accuracy of the questionnaire items (DeVellis & Thorpe, 2021). Based on the expert's guidance, those redundant, ambiguous or not closely related to the organizational identification of teachers in private higher education were revised or deleted. Subsequently, the questionnaire was piloted on 51 teachers to test the reliability and validity of the instrument, and to provide a basis for further improvement. For the first part of the questionnaire measuring organizational identification, the pretest indicators, including item discrimination, internal consistency, and exploratory factor analysis, were generally satisfactory, although several items presented problems and required revision or deletion. For example, item 3 was deleted due to poor item discrimination, and three items (9, 10, 11) were removed due to factor misloading or cross-loadings. A summary of the reliability and validity results is presented in Table 2.

Table 2. Summary of Reliability and Validity Analysis

Factor	Item	Factor loading	Communality	Cronbach's α	KMO value	Bartlett's test	Total variance explained
Existence identification	1	0.846	0.721	0.872	0.858	$\chi^2=796.556$; $p < 0.01$	71.511%
	2	0.935	0.899				
	4	0.742	0.800				
	5	0.614	0.711				
Belongingness identification	6	0.805	0.735	0.921			
	7	0.650	0.672				
	8	0.738	0.692				
	9	0.477	0.704				
	10	0.504	0.695				
	11	0.516	0.607				
	12	0.639	0.702				
	13	0.662	0.611				
Success identification	14	0.635	0.667	0.917			
	15	0.764	0.652				
	16	0.710	0.655				
	17	0.806	0.732				
	18	0.875	0.795				
	19	0.819	0.729				
	20	0.742	0.807				

Note: Exploratory factor analysis was conducted using principal component analysis with Varimax rotation.

After revision, the formal questionnaire showed high reliability coefficients (the overall reliability coefficient for organizational identification was 0.938, and those for the three dimensions were 0.814, 0.904, and 0.886). Each dimension of teachers' organizational identification showed a strong correlation with the total score ($p < 0.01$), and the dimensions were also substantially correlated with one another, as shown in Table 3.

Table 3. Pearson correlations of organizational identification dimensions (N = 268)

	Total	Existence identification	Belongingness identification	Success identification

Total	—			
Existence identification	0.825 ($p<0.001$)	—		
Belongingness identification	0.925 ($p<0.001$)	0.644 ($p<0.001$)	—	
Success identification	0.919 ($p<0.001$)	0.623 ($p<0.001$)	0.801 ($p<0.001$)	—

The final formal questionnaire retained 16 items in the first part (teachers' organizational identification), including 4 items for existence identification, 6 items for belongingness identification, and 6 items for success identification. All items adopted a five-point Likert scale (1 = "strongly disagree" to 5 = "strongly agree"), with higher scores indicating stronger organizational identification. The second part of the questionnaire consisted of 7 items on basic information, including gender, age, educational level, academic title, full-time status, administrative duty, and establishment status. Taken together, the final questionnaire contained 23 items.

3.1.2. Data Collection

The formal questionnaire was administered through the Wenjuanxing online survey platform to teachers from private colleges and universities in Shanghai, using a convenience sampling approach. The survey link was distributed via WeChat in May 2025. To protect anonymity and personal privacy, respondents were invited, but not required, to indicate their institution. Based on the available responses, teachers from 14 private higher education institutions in Shanghai were identified. A total of 280 questionnaires were collected, of which 268 were valid, yielding a valid response rate of 95.7%. The sample information is presented in Table 4. As institutional identification was optional, participation by individual institutions is not reported.

Table 4. Descriptive Statistics of the Survey Sample (N = 268)

Basic information	Category	Count	Percentage
Gender	Male	52	19.4%
	Female	216	80.6%
Age	30 and below	96	35.8%
	31—40	122	45.5%
	41—50	40	14.9%
	51 and above	10	3.7%
Educational level	Bachelor and below	37	13.8%
	Master	187	69.8%
Academic title	Doctor	44	16.4%
	Assistant and below	108	40.3%
	Lecturer	116	43.3%
	Associate Professor	41	15.3%
Full-time status	Yes	234	87.3%
	No	34	12.7%
Administrative duty	Yes	115	42.9%
	No	153	57.1%
Establishment status	Yes	23	8.6%
	No	245	91.4%

3.1.3. Data Analysis

SPSS was used to analyze the data from the questionnaire survey. Descriptive statistics (means and standard deviations) were calculated for the overall organizational identification score and for each of

the three dimensions. Independent-samples t-tests were conducted to compare organizational identification across binary groups (e.g., gender, full-time status, administrative duty, establishment status). One-way ANOVA was used to test group differences for variables with more than two categories (e.g., age groups, educational level, academic title). When the omnibus ANOVA indicated a statistically significant difference, post-hoc pairwise comparisons were conducted using the Games-Howell procedure, which does not assume equal variances and is suitable for unequal sample sizes. Furthermore, to address the risk of Type I error inflation from multiple comparisons (28 tests in total), the Benjamini-Hochberg false discovery rate (FDR) correction procedure was applied using a false discovery rate of 0.05.

3.2 Semi-structured Interviews

3.2.1. Interview Protocol

The interview protocol was structured into three sections with 9 open-ended questions. The first gathered background details to contextualize participants' current experiences and inform later strategic analysis. The second explored their perceptions of the organization. The third elicited suggestions for institutional improvements. After drafting the protocol, two teachers from private higher education were selected for pilot interviews, and the questions and wording were further revised and refined.

3.2.2. Data Collection

All interviewees were teachers from private higher education in Shanghai. Through purposive sampling with snowball recruitment, the scope of interviewees was expanded as much as possible across different institutions and years of service. Based on the principle of data saturation, although relatively sufficient material had been obtained by the time eight interviews had been completed, interviews with three additional teachers were conducted to further verify and supplement the information gathered. Basic information on the interviewees is shown in Table 5.

Table 5. Basic Information of Interviewees

No.	Gender	Educational level	Years of service (approx., year)
1	Male	Master	3
2	Female	Master	4
3	Male	Master	3
4	Female	Doctor	11
5	Female	Master	2
6	Male	Doctor	2
7	Male	Master	1.5
8	Male	Master	7
9	Female	Master	0.5
10	Female	Master	5
11	Female	Master	6

According to the interviewees' preferences, telephone interviews were conducted in June 2025. Original materials were collected by combining note-taking with audio recording. Before each interview, full communication was conducted to explain the research purpose and obtain informed consent. It was promised that no personal privacy would be involved and that all data would be kept confidential to ensure the richness and authenticity of the information obtained. Interviewees are referred to in the text as Teacher followed by a number (such as Teacher 1, Teacher 2), consistent with the identifiers in Table 5, and real names are not used.

3.2.3. Data Analysis

Within twenty-four hours after each interview, the audio recordings were transcribed verbatim, and the transcripts were organized and archived in NVivo. The analysis followed a conventional content analysis approach (Hsieh & Shannon, 2005). During data processing, the recordings were repeatedly reviewed, and interpretations were checked and refined against the transcripts to ensure accuracy in the analysis of the textual materials. The primary researcher conducted the coding and identified themes from the interview data. To enhance accuracy, the coding process was carried out in two rounds. The preliminary interpretations were then discussed among the research team to reach agreement. Participants were not asked to review or comment on the findings. In addition, the interview excerpts cited in the findings were translated into English from the original Chinese to facilitate reporting. Efforts were made to preserve the participants' intended meanings during translation, and the translated excerpts were reviewed by a bilingual researcher for accuracy.

4. Findings

4.1. Current Status of Teachers' Organizational Identification

According to the high or low scores obtained in each part of the questionnaire, the corresponding strength of teachers' organizational identification can be indicated. The mean score of organizational identification was obtained by summing the item scores and dividing by the number of items, which allows understanding of the overall situation of organizational identification and each dimension.

4.1.1 Overall Level of Organizational Identification

Organizational identification reflects teachers' psychological perceptions of the institution. By calculating the mean scores, the overall level of organizational identification among teachers in private higher education can be understood, and the identification status for each dimension can also be further examined. The specific scores are shown in Table 6.

Table 6. Descriptive Statistics of Teachers' Organizational Identification (N = 268)

Dimension	Organizational identification (total)	Existence identification	Belongingness identification	Success identification
Mean	3.517	2.759	3.878	3.661
Standard deviation	0.740	0.943	0.788	0.789

Table 6 reveals a mean organizational identification score of 3.517, slightly above the midpoint of the five-point Likert scale, suggesting a moderate overall level. The mean score of belongingness identification is the highest at 3.878, which is higher than the overall mean. This indicates that teachers show a relatively strong sense of belongingness identification with private colleges and universities. However, the mean of belongingness identification does not reach 4.000, which indicates that the perceived belongingness identification is still not strong enough. The mean score of success identification is 3.661, which is also higher than the overall mean and falls in the upper-middle range, indicating that the perception of success identification is not ideal either. In addition, existence identification has the lowest mean score at 2.759, which is far below the overall mean. This indicates that the low level of existence identification is a primary factor leading to the overall level of organizational identification not being high in private higher education.

To understand the situation in more detail, the mean scores of the items under each dimension are examined separately. Where appropriate, qualitative interview data are used to contextualize and explain specific quantitative patterns.

4.1.2. Existence Identification

Existence identification refers to teachers' psychological perception that the institution provides material support to meet basic livelihood needs. For teachers in private higher education, establishment status influences income to some extent and may also affect perceived job stability, which in turn shapes organizational identification. Therefore, existence identification mainly involves perceived job stability and compensation. Job stability covers two items, Q1 and Q2, and compensation covers two items, Q3 and Q4. The scores are shown in Table 7.

Table 7. Descriptive Statistics for Items of Existence Identification (N = 268)

Item	Mean	Standard deviation
Q1	2.28	1.214
Q2	3.00	1.275
Q3	2.78	1.098
Q4	2.98	1.110

As shown in Table 7, Q1 has the lowest mean score at 2.28. This indicates that teachers care about formal establishment status, and the fact that most teachers in private higher education do not have it weakens their organizational identification. Interviews show that teachers may consider establishment status unimportant before entry or shortly after entry, but its salience tends to increase as they work longer and understand the system more fully. Some focus on the benefits associated with establishment status, as noted by one teacher:

“Our housing fund and social insurance cannot compare with those of teachers on establishment. Private colleges are indeed lacking in this respect, which is exactly why they have trouble retaining teachers. Young teachers especially face survival pressure.” (Teacher 3)

Others focus more on stability:

“I used to think establishment status was not important, but now I really think it matters. I feel it brings a stronger sense of security and belonging.” (Teacher 5)

Although Q2 is at a general level in absolute terms, it is the highest within this dimension. This suggests that the labor contract partly meets teachers' need for stability and thus raises organizational identification. As one teacher stated:

“I do not have establishment status, but I do not think stability is a big problem. As long as you do not make serious mistakes, the school will not easily dismiss a teacher. The labor contract law is there. I do not have many concerns about stability.” (Teacher 11)

The mean scores of Q3 and Q4 are both below 3.00. This indicates dissatisfaction with pay and benefits, which reduces organizational identification. Almost all interviewees expressed dissatisfaction with compensation or benefits:

“Salaries of teachers in private colleges and universities are far behind public ones. Our workload is two or three times that of public universities, while our pay is about one half or one third.” (Teacher 7)

“The salary and benefits of teachers at private colleges and universities are low. For locals

like me in Shanghai, after daily expenses there is little left. If you have a family and rely on this single income, the pressure is heavy. For non-locals, with rent or a mortgage, the pressure is even greater.” (Teacher 8)

4.1.3. Belongingness identification

Belongingness identification refers to the psychological perception formed in interpersonal interaction and affective exchange. For teachers in private higher education, belongingness identification arises partly from the institution-level experiences and partly from colleagues. The institution-related aspect includes three items (Q5, Q6, Q7). The colleague-related aspect includes three items (Q8, Q9, Q10). The scores are shown in Table 8.

Table 8. Descriptive Statistics for Items of Belongingness Identification (N = 268)

Item	Mean	Standard deviation
Q5	3.68	0.975
Q6	3.73	1.004
Q7	3.82	1.000
Q8	3.81	1.043
Q9	4.29	0.759
Q10	3.93	0.941

As shown in Table 8, the three items under the colleague-related aspect all have mean scores above 3.80, indicating a relatively high level of identification. Q9 has the highest mean at 4.29, which exceeds the fourth level on the five-point Likert scale and is clearly higher than the other items. This suggests that a strong sense of belongingness identification mainly comes from good collegial relations. As one teacher noted:

“We are not just colleagues; we chat, and sometimes go out together. I think our relationships are pretty good.” (Teacher 9)

The mean score of Q10 is the second highest at 3.93, which indicates that teachers not only get along well but also find like-minded partners for collaboration based on interests or development goals. This supports stronger identification and benefits both individual and group development. As one teacher observed:

“There is competition among young teachers, but there is more cooperation. We encourage one another to apply for projects together. Senior teachers also care about young teachers’ growth, offering help to first-time project members.” (Teacher 3)

Within this dimension, Q8 has the lowest mean, indicating that relationships with leaders are not as close as relationships among peers and that leaders need to show more care and support. As one teacher put it:

“Leaders should think more from teachers’ perspective, whether about professional planning and development or about compensation and welfare benefits. They should think more for us, rather than considering things from a profit-oriented angle. It can be disheartening.” (Teacher 5)

For the institution-related aspect, all three means are above 3.60, which indicates an upper-middle level of identification. Q7 has the highest mean at 3.82, suggesting general satisfaction with the overall environment. As some teachers emphasized:

“Our working environment is quite good and humane. In private colleges, if implementation is too rigid without considering teachers’ feelings, many will leave. Because the environment is comfortable, and there are not many additional demands, teachers just want to do their work well.” (Teacher 1)

“The cultural and learning atmosphere is good among both students and teachers. As young teachers, we have a strong desire to develop, and the environment naturally motivates us. My inertia is worn down and my drive is stimulated, so I maintain an upward state in both work and study.” (Teacher 3)

The means of Q5 and Q6 are relatively lower within this dimension, which indicates that teachers’ willingness to remain in the same school for a long time is not very strong. As one teacher commented:

“I heard that half of the faculty in a certain college left. How can the college keep running? If teachers feel uncomfortable at work, face heavy pressure, and receive poor compensation, many will leave if they have better opportunities.” (Teacher 7)

4.1.4. Success Identification

Success identification refers to the psychological perception formed as teachers develop and improve themselves with support from the institution. For teachers in private higher education, it is mainly influenced by institution resource support and personal development. The resource-related aspect includes Q11, Q12, and Q13. The personal-development aspect includes Q14, Q15, and Q16. The scores are shown in Table 9.

Table 9. Descriptive Statistics for Items of Success Identification (N = 268)

Item	Mean	Standard deviation
Q11	3.61	1.073
Q12	3.38	1.103
Q13	3.38	1.041
Q14	3.96	0.858
Q15	3.84	0.907
Q16	3.79	0.925

As shown in Table 9, the highest mean among the three resource-related items is 3.61, which is lower than the lowest mean among the three personal-development items (3.79). This suggests that teachers generally acknowledge that the institution facilitates personal development, but limited resources constrain development to some extent and weaken success identification. Q11 has a mean of 3.61, indicating an upper-general level. It shows that the institution provides some training opportunities, though issues remain with content and delivery. As one teacher noted:

“The institution offers quite a few training opportunities, especially when we first join, but the quality is often not ideal. The better programs usually have very few places and tend to go to middle managers or to those with longer tenure, so junior teachers have limited access.” (Teacher 2)

Q12 and Q13 share the same mean of 3.38, which is relatively low. This indicates shortages in electronic and library resources and in collaboration with external professional bodies. Examples shared by teachers include:

“Our library holdings are limited, and electronic databases are not as rich. We still do not subscribe to CNKI (a Chinese academic literature database) due to cost, which creates problems for literature retrieval in research and teaching.” (Teacher 3)

“Cooperation with enterprises is insufficient. If strengthened, it would greatly help both our discipline and students’ development.” (Teacher 8)

Within the personal-development aspect, Q14 has the highest mean at 3.96, indicating considerable alignment between teachers’ and the school’s development goals. As one teacher remarked:

“Our program is new and challenging, but the school has provided strong support, from funding to project applications, and that has helped us grow. I am very grateful for the school’s attention to our development.” (Teacher 3)

The means of Q15 and Q16 are around 3.80, showing that teachers recognize the value of their work for applying skills and for personal growth, while also facing obstacles:

“Private higher education provides a relatively free platform. For practice-based creative work, the school supports us with venues, equipment, studios, and personnel, and there is flexibility in what we create.” (Teacher 9)

“Young teachers often take on administrative tasks. Teaching already brings heavy pressure, and additional administrative duties increase the burden and hinder development.” (Teacher 2)

4.2. Differences in Organizational Identification

4.2.1. Differences by Gender

Using gender as the grouping variable, organizational identification and its three dimensions as the test variables, independent-samples t tests were conducted. The results are shown in Table 10.

Table 10. Differences in Organizational Identification by Gender (N = 268)

	Gender	Mean	SD	<i>t</i>	<i>P</i>	Cohen’s <i>d</i>	95% CI
Organizational identification (total)	Male	3.546	0.749	0.311	0.756	0.048	[-0.194, 0.265]
	Female	3.510	0.739				
Existence identification	Male	2.813	1.020	0.452	0.651	0.070	[-0.243, 0.375]
	Female	2.747	0.925				
Belongingness identification	Male	3.869	0.820	-0.097	0.923	0.015	[-0.262, 0.238]
	Female	3.880	0.782				
Success identification	Male	3.711	0.804	0.513	0.608	0.079	[-0.184, 0.309]
	Female	3.649	0.787				

Note: * $p < 0.05$, ** $p < 0.01$

No significant differences were found between male and female teachers on the total score of organizational identification or on any of the three dimensions. This suggests that male and female

teachers in private higher education exhibit similar levels of organizational identification. The result is consistent with findings reported by Zhou (2018) on university teachers' organizational identification.

4.2.2. Differences by Age

Using age as the grouping variable, one-way ANOVA was conducted on the total score of organizational identification and its three dimensions. The results are shown in Table 11.

Table 11. Differences in Organizational Identification by Age (N = 268)

	Age	Mean	SD	95% CI	F	p	η^2
Organizational identification (total)	30 and below	3.516	0.692	[3.376, 3.656]	2.407	0.068	0.027
	31—40	3.515	0.722	[3.386, 3.644]			
	41—50	3.384	0.830	[3.119, 3.649]			
	51and above	4.081	0.871	[3.449, 4.713]			
Existence identification	30 and below	2.849	0.905	[2.665, 3.033]	2.707	0.046*	0.030
	31—40	2.674	0.910	[2.511, 2.837]			
	41—50	2.631	1.000	[2.312, 2.950]			
	51and above	3.450	1.212	[2.570, 4.330]			
Belongingness identification	30 and below	3.873	0.766	[3.718, 4.028]	1.894	0.131	0.021
	31—40	3.908	0.739	[3.776, 4.040]			
	41—50	3.688	0.909	[3.399, 3.977]			
	51and above	4.317	0.954	[3.620, 5.014]			
Success identification	30 and below	3.604	0.730	[3.456, 3.752]	2.321	0.076	0.026
	31—40	3.682	0.801	[3.539, 3.825]			
	41—50	3.583	0.858	[3.310, 3.856]			
	51and above	4.267	0.746	[3.732, 4.802]			

Note: * $p < 0.05$, ** $p < 0.01$

As indicated in Table 11, the total score of organizational identification does not differ significantly across age groups. At the dimensional level, only existence identification shows a significant difference at the 0.05 level, although the effect size was small ($\eta^2 = 0.030$). Comparison of means suggests that teachers aged 51 and above report notably higher existence identification than teachers in the other age groups. Post-hoc comparisons using the Games-Howell procedure further indicated that teachers aged 51 and above reported significantly higher existence identification than those aged 31–40 ($p = 0.032$) and 41–50 ($p = 0.043$). This pattern may stem from older teachers' longer tenure or senior hires as experts or professors, which often come with greater job security and better pay. For the other age groups, the mean scores of existence identification are all below 3.00, indicating a level lower than general agreement. This pattern may reflect the absence of establishment status for most teachers in private higher education and comparatively lower pay and benefits, which in turn weakens existence identification.

4.2.3. Differences by Educational Level

Using educational level as the grouping variable, one-way ANOVA was conducted on the total score of organizational identification and its three dimensions. The results are shown in Table 12.

Table 12. Differences in Organizational Identification by Educational Level (N = 268)

	Educational level	Mean	SD	95% CI	F	p	η^2
Organizational identification (total)	Bachelor and below	3.963	0.647	[3.748, 4.178]	9.405	0.000**	0.066
	Master	3.411	0.735	[3.305, 3.517]			
	Doctor	3.591	0.697	[3.380, 3.802]			
	Bachelor and below	3.358	0.906	[3.057, 3.659]	9.905	0.000**	0.070

Existence identification	Master	2.630	0.934	[2.495, 2.765]			
	Doctor	2.807	0.823	[2.557, 3.057]			
Belongingness identification	Bachelor and below	4.248	0.693	[4.018, 4.478]	6.220	0.002**	0.045
	Master	3.779	0.790	[3.665, 3.893]			
	Doctor	3.989	0.759	[3.759, 4.219]			
Success identification	Bachelor and below	4.081	0.749	[3.833, 4.329]	7.036	0.001**	0.050
	Master	3.565	0.768	[3.454, 3.676]			
	Doctor	3.716	0.807	[3.471, 3.961]			

Note: * $p < 0.05$, ** $p < 0.01$

As indicated in Table 12, the total score and all three dimensions differ significantly across degree groups at the 0.01 level, with small to medium effect sizes ($\eta^2 = 0.045$ -0.070). Post-hoc comparisons further indicated that teachers with a bachelor's degree or below scored significantly higher than those with a master's degree on the total score and all three dimensions (all $p < 0.01$). Significant differences were also observed between teachers with a bachelor's degree or below and those with a doctoral degree for the total score ($p = 0.033$) and existence identification ($p = 0.006$), while no significant differences were found for the other dimensions.

Mean comparisons also show consistently higher scores among teachers with a bachelor's degree or below. One plausible explanation is that, as higher education has expanded in China and the supply of master's and doctoral graduates has increased, private higher education have raised hiring requirements. For teachers with only a bachelor's degree or below, obtaining a college position may already meet expectations, so their demands on the institution are lower and identification is correspondingly higher. Although no statistically significant differences were observed between teachers with doctoral and master's degrees, mean scores were slightly higher among doctorates across the total and all dimensions. This pattern may reflect the relatively scarce supply of doctoral holders in private higher education (16.4% in this study) and the more favorable employment conditions often offered to attract them.

4.2.4. Differences by Academic Title

Using academic title as the grouping variable, one-way ANOVA was conducted on the organizational identification total score and its three dimensions. The results are shown in Table 13.

Table 13. Differences in Organizational Identification by Academic Title (N = 268)

	Academic title	Mean	SD	95% CI	F	p	η^2
Organizational identification (total)	Assistant and below	3.595	0.679	[3.466, 3.724]	3.753	0.011*	0.041
	Lecturer	3.387	0.724	[3.254, 3.520]			
	Associate Professor	3.607	0.872	[3.333, 3.881]			
	Professor	4.521	0.036	[4.428, 4.614]			
Existence identification	Assistant and below	2.917	0.909	[2.744, 3.090]	4.290	0.006**	0.046
	Lecturer	2.580	0.863	[2.421, 2.739]			
	Associate Professor	2.762	1.140	[2.403, 3.121]			
	Professor	4.000	0.250	[3.472, 4.528]			
Belongingness identification	Assistant and below	3.929	0.759	[3.785, 4.073]	2.607	0.052	0.029
	Lecturer	3.778	0.781	[3.634, 3.920]			
	Associate Professor	3.955	0.853	[3.687, 4.223]			
	Professor	4.889	0.192	[4.472, 5.306]			
Success identification	Assistant and below	3.713	0.721	[3.576, 3.850]	2.904	0.035*	0.032
	Lecturer	3.534	0.792	[3.389, 3.679]			
	Associate Professor	3.821	0.910	[3.535, 4.107]			
	Professor	4.500	0.167	[4.072, 4.928]			

Note: * $p < 0.05$, ** $p < 0.01$

As indicated in Table 13, except for belongingness identification, the organizational identification total score and the other two dimensions show significant differences across academic titles. However, the effect sizes ($\eta^2 = 0.029-0.046$) suggest that the observed differences are statistically small. Post-hoc comparisons indicated that professors reported significantly higher organizational identification than the other academic title groups on the total score and existence identification (all $p < 0.05$). Mean comparisons further show that professors report the strongest organizational identification, with mean scores above 4.00, indicating a high level of identification. By contrast, teachers with the lecturer title generally report the lowest means on the organizational identification total and all dimensions. A significant difference for success identification was also observed only between professors and lecturers ($p = 0.020$). This likely reflects that lecturers tend to be at a relatively lower level in terms of salary and perceived recognition, and face greater difficulty in promotion, which may depress their organizational identification at this stage. This pattern is consistent with career-stage perspectives, which suggest that early-career teachers often face greater uncertainty and weaker organizational attachment (Bennett et al., 2018).

4.2.5. Differences by Full-Time Status

Using full-time status as the grouping variable, independent-samples *t* tests were conducted on the organizational identification total score and its three dimensions. The results are reported in Table 14.

Table 14. Differences in Organizational Identification by Full-time Status (N = 268)

	Full-time status	Mean	SD	95% CI	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Organizational identification (total)	Yes	3.511	0.747	[3.415, 3.607]	-0.321	0.748	0.059
	No	3.555	0.696	[3.314, 3.796]			
Existence identification	Yes	2.740	0.957	[2.617, 2.863]	-0.862	0.389	0.158
	No	2.890	0.840	[2.599, 3.181]			
Belongingness identification	Yes	3.881	0.786	[3.780, 3.982]	0.160	0.873	0.029
	No	3.858	0.815	[3.576, 4.140]			
Success identification	Yes	3.656	0.790	[3.555, 3.757]	0.276	0.782	0.051
	No	3.696	0.793	[3.422, 3.970]			

Note: * $p < 0.05$, ** $p < 0.01$

As shown in Table 14, no significant differences were found on the total score or on any dimension, indicating that full-time and part-time teachers in private higher education exhibit similar levels of organizational identification.

4.2.6. Differences by Administrative Duty

Using whether a teacher holds an administrative duty as the grouping variable, independent-samples *t* tests were performed on the total score of organizational identification and its three dimensions. The results are shown in Table 15.

Table 15. Differences in Organizational Identification by Administrative Duty (N = 268)

	Administrative duty	Mean	SD	95% CI	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Organizational identification (total)	Yes	3.550	0.661	[3.428, 3.672]	0.649	0.517	0.078
	No	3.492	0.795	[3.365, 3.619]			
Existence identification	Yes	2.744	0.895	[2.578, 2.908]	-0.238	0.812	0.029
	No	2.771	0.980	[2.615, 2.927]			
Belongingness identification	Yes	3.933	0.703	[3.804, 4.062]	1.021	0.308	0.123
	No	3.837	0.846	[3.702, 3.972]			

Success identification	Yes	3.704	0.736	[3.568, 3.840]	0.791	0.430	0.096
	No	3.629	0.828	[3.497, 3.761]			

Note: * $p < 0.05$, ** $p < 0.01$

As shown in Table 15, no significant differences emerged for the total score or for any dimension. Holding an administrative duty does not appear to have a significant effect on organizational identification among teachers in private higher education.

4.2.7. Differences by Establishment Status

Using establishment status as the grouping variable, independent-samples t tests were conducted on the total score of organizational identification and its three dimensions. The results are shown in Table 16.

Table 16. Differences in Organizational Identification by Establishment Status (N = 268)

	Establishment status	Mean	SD	95% CI	t	p	Cohen's d
Organizational identification (total)	Yes	3.807	0.776	[3.472, 4.142]	1.978	0.049*	0.431
	No	3.490	0.732	[3.398, 3.582]			
Existence identification	Yes	2.990	1.029	[2.546, 3.432]	1.224	0.222	0.267
	No	2.738	0.934	[2.621, 2.855]			
Belongingness identification	Yes	4.152	0.767	[3.819, 4.485]	1.752	0.081	0.382
	No	3.852	0.786	[3.753, 3.951]			
Success identification	Yes	4.007	0.837	[3.650, 4.364]	2.217	0.027*	0.483
	No	3.629	0.778	[3.531, 3.727]			

Note: * $p < 0.05$, ** $p < 0.01$

As shown in Table 16, the total score of organizational identification and success identification differ significantly at the 0.05 level ($d = 0.431$ and 0.483), with teachers having establishment scoring higher than those without. Although existence identification and belongingness identification do not reach statistical significance, the means are also higher for the establishment group, suggesting generally stronger organizational identification.

4.2.8. Multiple Comparisons Correction

Given the relatively large number of group difference tests conducted in this study, a multiple comparisons correction was applied to control the risk of increased Type I errors. Specifically, the Benjamini–Hochberg false discovery rate (FDR) procedure with a false discovery rate of 0.05 was used in all statistical tests. This approach can control false positives while maintaining statistical power in studies involving multiple comparisons.

Table 17. Summary of Results Before and After Benjamini-Hochberg FDR Correction

Variable	Dimension	Uncorrected p	FDR-adjusted p	Significant after FDR	Final Classification
Education level	Overall identification	0.000	0.001	Yes	Significant
Education level	Existence identification	0.000	0.001	Yes	Significant
Education level	Belongingness identification	0.002	0.014	Yes	Significant
Education level	Success identification	0.001	0.009	Yes	Significant
Academic title	Existence identification	0.006	0.033	Yes	Significant
Academic title	Overall identification	0.011	0.051	No	Exploratory trend
Academic title	Success identification	0.035	0.122	No	Exploratory trend
Establishment status	Overall identification	0.049	0.132	No	Exploratory trend

Establishment status	Success identification	0.027	0.108	No	Exploratory trend
Age	Existence identification	0.046	0.132	No	Exploratory trend

Note: The false discovery rate was set at 0.05. Only tests that were originally significant or marginally significant are presented. All other tests were non-significant before and after correction.

The results of the FDR adjustment are presented in Table 17. After correction, the differences associated with educational level remained statistically significant for the total organizational identification score and all three dimensions. In contrast, only the difference in existence identification across academic titles remained significant, whereas the differences in total and success identification by academic title were no longer statistically significant. In addition, the initially significant differences related to establishment status and age did not remain significant after the FDR adjustment and are therefore treated as exploratory trends rather than confirmed group differences. These findings should be interpreted with caution in practice, as they may reflect unstable patterns that require further verification in studies with larger samples.

5. Discussion

5.1 Discussion of the Findings

The results of this study offer a nuanced conceptual understanding of organizational identification among teachers in private higher education. Overall, the survey indicates that teachers in private higher education exhibit only moderate organizational identification. This level appears lower than those reported in some recent studies of teachers' organizational identification in higher education settings (Cui & Suntrayuth, 2023; Khanal & Arora, 2024), suggesting that institutional conditions in private higher education may pose additional challenges for sustaining strong identification. Across dimensions, the levels are unbalanced. The mean level of existence identification is too low, which is the principal reason for the overall level being modest. Although the indicators of belonging and success are relatively high, both scores are below 4.00, indicating that the identification in these two dimensions is not particularly strong. These findings not only indicate that the teachers of private higher education lack a sufficient sense of organizational identification, but also reveal numerous practical difficulties existing in the school's organizational management.

Early organizational identification research usually focuses on psychological belonging or value alignment (Ashforth & Mael, 1989; Christopher & Divya, 2024). However, in the special context of private higher education, this study found that existence identification is not only the foundation of the identification structure but also a key condition shaping the overall sense of identification. Teachers are usually dissatisfied with salary and welfare benefits and attach considerable importance to establishment status, which they view as bringing better benefits and job stability. Although the signing of labor contracts partially compensates for the pursuit of stability, other deficiencies still require sufficient attention from school administrators. Especially when teachers encounter obstacles in their pursuit of career achievements, they will exhibit extreme sensitivity to their survival needs. This phenomenon is consistent with the frustration-regression mechanism of the ERG theory, suggesting that when achievement needs are frustrated, teachers may shift their attention toward survival and belonging concerns. (Alderfer, 1972). This means that organizational identification is not a static psychological state of teachers, but rather a dynamic process of constant negotiation between survival guarantees and self-actualization.

Although the existence identification was at a low point, belongingness identification received the highest score. This coexistence of high belongingness identification and low existence identification reveals a unique emotional compensation mechanism within private higher education. A similar pattern has been observed among university teachers, where strong collegial support helps maintain organizational identification even when material rewards and job stability are limited (Chen et al., 2025).

In this study, belongingness identification appears to be closely associated with positive collegial relationships. Teachers get along well and can find like-minded partners for collaboration based on interests or development goals. When institutional support is insufficient, teachers tend to seek informal support within small groups. However, leaders' care and support for teachers are perceived as insufficient. This localized form of identification may help maintain organizational functioning, but if this horizontal colleague identification fails to transform into vertical organizational trust, it is unlikely to develop into a stronger, institution-wide identification. This explains another finding of this study: while teachers are generally satisfied with the overall institutional environment, their willingness to remain at the same institution for a long time is not strong. The persistence of this phenomenon over time may easily lead to the formation of subgroup cultures, thereby weakening the organizational cohesion and long-term collective commitment (Rovetta et al., 2025).

For success identification, the results show that although the mean value is higher than that of existence identification, it still does not reach the level of strong identification. This indicates a form of vision mismatch between the teachers' professional growth and institutions' goals. While teachers believe that their institutions provide some training opportunities, problems exist in the content and delivery of these programs. Shortages are also reported in library and electronic resources and in collaboration with external professional organizations. Although teachers tend to agree with the institutional development goals and with the value of their work, the exercise and development of personal capabilities are still constrained in several ways. This lack of professional satisfaction may hinder teachers' long-term commitment and reduce institutions' ability to retain staff with academic ambitions (Wu, 2024). From this perspective, success identification should not be reduced to a matter of short-term performance incentives, but rather understood as a form of value-based alignment between institutional leadership and teachers.

In addition to the imbalance across identification dimensions, the observed differences among teacher groups may reflect the uneven distribution of institutional resources within private higher education. According to the difference analyses, no significant differences appear by gender, full-time status, or administrative duty. Teachers aged 51 and above tended to report higher existence identification than other age groups. Teachers with a bachelor's degree or below score significantly higher than those with a master's degree on organizational identification and its dimensions, while teachers with a doctoral degree generally report slightly higher scores than those with a master's degree. Teachers with the professor title show stronger existence identification than other teachers and tend to report higher scores on the other dimensions, while those with the lecturer title report the lowest mean scores on the total and across dimensions. Teachers with establishment status tended to report higher success identification than those without, and they also showed relatively higher scores on existence and success identification overall.

These patterns suggest that uniform management approaches may be insufficient. Instead, institutional leaders should consider differentiated strategies that address the distinct needs and identification profiles of diverse teacher groups. It should also be noted that several statistically significant differences in this study were associated with relatively small effect sizes ($\eta^2 < 0.05$), indicating that the magnitude of these differences was limited. In this sense, institutional responses may be better directed toward targeted support measures rather than broad structural adjustments.

In conclusion, institutional leaders should recognize that the generally moderate level of teachers' organizational identification is a signal of long-term strategic risk for the school. The differentiation of teachers' organizational identification can serve as a starting point for targeted governance intervention measures. Enhancing overall identification should not rely on symbolic appeals, but should be regarded as a systematic governance process that aligns institutional structures, leadership practices, and teachers' professional trajectories.

5.2. Implications

Based on the above findings, this study proposes the following governance and leadership implications to enhance teachers' organizational identification in private higher education institutions.

5.2.1. Implementing Coordinated Strategies to Enhance Existence Identification

From a strategic management perspective, institutional leaders must prioritize the stabilization of the teachers' existence needs as a prerequisite for long-term governance (Zhan et al., 2023; Xu & Pang, 2024). Although it is difficult in the short term to raise salaries in private colleges and universities to an ideal level and establishment status is hard to resolve, leadership can still improve existence identification through refined administrative mechanisms.

First, optimize performance leadership and appraisal systems. To effectively motivate teachers' work and development, leadership should build more rational and systematic performance appraisal systems that truly guide and incentivize. When teachers' work is reasonably quantified and appropriately rewarded, income will naturally increase and motivation for work and development will be strengthened. Performance indicators for teaching, research, and social service should be refined, tailored to the characteristics of faculties and even specific programs, and avoid one-size-fits-all arrangements. If administrative duties cannot be avoided in the short run, performance rewards for teachers who shoulder such duties should be increased to prevent dissatisfaction caused by mismatched effort and return. In addition, young teachers deserve greater recognition and support. Although their salaries and hourly wages are often relatively low, they usually demonstrate strong initiative and ambition. Educational leaders should take these development characteristics into account. For instance, this goal can be achieved by increasing the research rewards for early career personnel. For senior teachers, a performance-based reward system can motivate them to play exemplary roles and make contributions in teaching, research, and social services. This not only increases their income but also enhances the overall quality of the faculty. In short, leaders should promptly improve the evaluation system and ensure that the rewards are commensurate with the efforts through effective institutional incentives.

Secondly, strengthening welfare management serves as a key means for stability. In the Chinese public sector, establishment status implies higher welfare benefits and greater stability. Therefore, leaders in the private sector should regard welfare benefits as a strategic tool to enhance their existence identification. The leadership should attach importance to holiday benefits, and more importantly, increase the contribution ratio to social insurance and housing provident funds, which are key components of employment benefits in the Chinese institutional context. Among them, the contribution ratio of housing provident fund is particularly worthy of attention, as it not only reflects the level of benefits but also directly affects housing pressure, thereby influencing teachers' perception of job stability. If stability linked to establishment status cannot be provided, the leadership should focus on addressing the housing issue. Increasing the contribution rate of housing provident funds or providing housing subsidies can effectively retain talents and ensure the healthy development of the teachers. For young teachers with accommodation needs, reasonable-priced and well-equipped on-campus housing or subsidized housing should be provided.

5.2.2. Building A Supportive Climate for Belongingness Identification

Educational leaders should act as the builders of the institutional culture, creating a positive organizational atmosphere as the foundation for belongingness. A strong sense of belonging can enable teachers to collaborate with each other and make progress together, which is a key competitive advantage in the management of private higher education. This is both the appeal of private higher education institutions and an aspect that can be effectively improved. To create such an atmosphere, the joint efforts of teachers and leaders are required.

First, it is necessary to lead the establishment of a harmonious and mutually supportive working environment. Despite numerous institutional flaws, many teachers can still find solace in the supportive and mutually helping relationships among their colleagues. Therefore, the leadership should cultivate a unique organizational culture, strengthen it as an advantage, attract loyal and cooperative teachers to join this team, and create a relaxed, harmonious and beneficial atmosphere through interaction, thereby stimulating the collective development momentum (Plaku et al., 2025). On one hand, institutions can regularly organize cultural and recreational activities and competitions to relieve work pressure and enhance teamwork spirit. On the other hand, the leadership should improve the office environment and add leisure facilities, so that teachers can maintain a good mental state and have convenient communication spaces. Cooperation should be encouraged, and senior teachers should play a guiding role in instructing new teachers, thereby inheriting the mutual assistance norms and forming a good faculty culture.

Second, build mutual trust between leaders and teachers. Leaders should extend more humanistic care so that teachers feel the warmth of the community, cultivate a sense of ownership, and become more willing to contribute while also finding satisfaction in the process. Care should not be limited to the work matters. Difficulties in life and family also require attention and timely assistance. Management should be flexible and people-oriented. It should respect the opinions of teachers, so that administrative work can serve teaching, research and social services, rather than being restricted or burdened merely for convenience. Fairness is important in aspects such as task allocation, promotion of positions, and distribution of training opportunities. The relevant processes should be objective, fair and transparent, and free from any favoritism. The institution should also establish an effective communication platform, such as holding regular reception days attended by senior leaders and teacher representatives, so that teachers can express their ideas, reduce communication barriers, and create a harmonious and equal atmosphere. Leaders should treat teachers of different titles and educational levels equally and with respect. Master's-degree holders and lecturers remain a backbone of institutional development; overemphasizing doctorates and professors while neglecting others can generate perceived inequity. Only when mutual trust is established can teachers fully integrate into the institutional community, cooperate effectively with leadership, and even participate in governance.

5.2.3. Expanding Resources to Boost Success Identification

Private colleges and universities are platforms for teachers' professional development. Institutional leadership should provide all available internal and external resources and opportunities for learning and collaboration. When teachers perceive strong institutional support for professional growth, success identification strengthens and developmental motivation increases.

First, align institutional resources with teachers' professional needs. Leadership should improve training quality by shifting from a quantity-driven to a quality-oriented developmental approach. Given the nature of academic work, professional development spans the entire career. With rapid knowledge renewal and social change, faculty must learn in a timely manner and participate in training. Since needs differ by discipline and career stage, training should reflect actual needs to encourage genuine participation. If leaders pursue training quotas while neglecting outcomes, resources will be wasted. Flexible training options should be provided so that teachers can choose programs according to their work schedules and needs. At the same time, infrastructure leadership must address hardware and e-resource constraints that hinder teachers' development. Teaching, research, and service equipment should be updated and supplemented in a timely manner to avoid hardware bottlenecks. With improved facilities, collaboration with external organizations and enterprises will also be strengthened. By facilitating access to needed e-resources, exemplified by national academic databases such as the China National Knowledge Infrastructure (CNKI) in China or equivalent scholarly resources in other systems, can help ease difficulties in literature retrieval.

Second, implement autonomous management and reasonable workload distribution. Highly educated staff tend to value autonomy, and whether this need is met affects the application of their abilities (Yang, 2013). Teachers in private higher education also hope to enjoy job autonomy comparable to that in public universities, so they can choose to prioritize teaching, research, or social service. Leaders should therefore adopt flexible and trust-based management and make teachers' work and development more convenient. For example, where appropriate, the fixed on-duty system can be adjusted so that teachers without classes are not required to remain in the office. To genuinely secure discretionary time, workloads should be lightened and responsibilities clearly defined. Many interviewees reported heavy teaching loads in private colleges and universities, often combined with administrative tasks, which consume excessive time and energy. To genuinely secure discretionary time for professional growth, leadership must actively recruit to expand the faculty, distribute teaching tasks reasonably, and define responsibilities clearly to avoid assigning unnecessary administrative duties that consume excessive energy.

In conclusion, the strategies proposed in this study constitute a governance oriented framework for private higher education in China. Aspects such as welfare enhancement through housing provident funds and specific performance systems are closely tied to Chinese policy and institutional structures. The broader ideas, such as fostering a supportive organizational climate, building leader-teacher trust, and aligning resources with professional development needs, may provide useful reference points for enhancing teachers' organizational identification among private higher education institutions in other countries. However, any transfer of these recommendations requires careful local adaptation. Differences in national policies, labor markets, funding mechanisms, and cultural expectations mean that direct application without contextual adjustment may limit effectiveness.

6. Conclusion

This study examined teachers' organizational identification in Chinese private higher education institutions using a mixed-methods approach grounded in ERG theory and the three-factor model. The quantitative findings described the overall status of teachers' organizational identification and identified differences across teacher groups, reflecting variations associated with institutional conditions. The qualitative findings provided contextual explanations for these patterns and deepened understanding of teachers' experiences within their institutional environment. Together, these findings contribute to a clearer understanding of how existence, belongingness, and success needs shape teachers' identification with their institutions in private higher education. The study also offers practical guidance for institutional leaders by outlining governance strategies to strengthen organizational support, build trust, and promote professional development.

While the study provides meaningful insights into teachers' organizational identification in private higher education, several limitations need to be acknowledged. First, the analysis in this study is mainly descriptive and comparative. Instead of modeling the interactions among multiple factors, the analysis focuses on the general trends and group variations in teachers' organizational identification. Despite being appropriate for the study's exploratory purpose, this study does not allow for the examination of causal relationships among the related factors and organizational identification. Multivariate or structural modeling techniques can be used in future studies to examine the relationships among these variables.

Second, the sampling strategy of this study mainly employs convenience sampling and snowball sampling. While this method helps to reach teachers from private higher education institutions, it also limits the representativeness of the sample. The participants are more likely to come from accessible networks and institutions that are more willing to participate in this research. This may limit the representation of teachers who are less accessible through institutional or professional networks. Future research can address this limitation by adopting a broader sampling framework, thereby enhancing the representativeness of the sample and supporting more reliable inferences.

Third, this study was conducted in Shanghai, one of the most economically developed cities in China. The local labor market is relatively active, and employment opportunities are abundant. Teachers in this context may have different expectations regarding job stability and organizational support. This may partly explain the relatively low level of existence identification observed in this study. The findings may not fully represent private colleges and universities in less affluent regions.

Fourth, another limitation concerns the qualitative component of the study. Although interview data was carefully coded and reviewed, formal inter-rater reliability statistics were not calculated. Future studies may strengthen methodological rigor by involving multiple coders and reporting inter-rater agreement to enhance the transparency and reliability of qualitative analysis.

Finally, this study focuses on China's private higher education institutions, which have distinctive features such as unique governance structures, employment arrangements, and regulatory environments. The private higher education in China therefore represents a specific analytical context, and the findings should be interpreted with caution when generalized to other countries or institutional settings. Comparative studies of different private higher education systems will help clarify which aspects of organizational identification are context-specific and which may be more broadly applicable.

References

- Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Performance*, 4(2), 142–175. [https://doi.org/10.1016/0030-5073\(69\)90004-X](https://doi.org/10.1016/0030-5073(69)90004-X)
- Alderfer, C. P. (1972). *Existence, relatedness, and growth: Human needs in organizational settings*. Free Press.
- Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14(1), 20–39. <https://doi.org/10.5465/amr.1989.4278999>
- Bennett, D., Roberts, L., Ananthram, S., & Broughton, M. (2018). What is required to develop career pathways for teaching academics? *Higher Education*, 75(2), 271–286. <https://doi.org/10.1007/s10734-017-0138-9>
- Borlaug, S. B., Tellmann, S. M., & Vabø, A. (2023). Nested identities and identification in higher education institutions—The role of organizational and academic identities. *Higher Education*, 85, 359–377. <https://doi.org/10.1007/s10734-022-00837-5>
- Buckner, E. (2025). Rethinking private higher education for changing realities. *International Higher Education*, (123), 51–52. <https://doi.org/10.6017/895b9e0d.59115bc9>
- Chen, K.-H., Hsieh, M. H. H., Maritz, A., & Shieh, C.-J. (2025). What affects organizational identification in universities? Examining the effects of psychological contract on organizational identification and commitment of teachers. *Acta Psychologica*, 261, 1–12. <https://doi.org/10.1016/j.actpsy.2025.105756>
- Christopher, B. P., & Divya, S. (2024). A decadal review of organizational identification: Insights from bibliometric analysis and content analysis. *Humanities and Social Sciences Communications*, 11(1), 1–20. <https://doi.org/10.1057/s41599-024-03990-7>
- Cui, J., & Suntrayuth, S. (2023). Perceived organization supports and organization identification on work outcomes of Chinese teachers in Thailand. *Asia Social Issues*, 16(6), 1–23. <https://doi.org/10.48048/asi.2023.259692>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications (5th ed.)*. SAGE Publications.
- Feng, B., Dou, G. J., & Zhan, X. Q. (2025). Negative workplace gossip and turnover intention among kindergarten teachers: psychological safety as a mediator and organizational identification as a moderator. *Frontiers in Psychology*, 16, 1588482. <https://doi.org/10.3389/fpsyg.2025.1588482>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Hu, X., & Ghasemy, M. (2024). The private undergraduate colleges in China: The development, challenges and opportunities. *Malaysian Journal of Social Sciences and Humanities*, 9(12), e003126. <https://doi.org/10.47405/mjssh.v9i12.3126>

Jun, K., Hu, Z., & Sun, Y. (2023). Impact of authentic leadership on employee turnover intention: Perceived supervisor support as mediator and organizational identification as moderator. *Frontiers in Psychology, 14*, 01-15. <https://doi.org/10.3389/fpsyg.2023.1009639>

Khanal, K., & Arora, N. (2024). Business school faculty perceptions of CSR initiatives on college/university organizational citizenship behavior: the mediating effect of organizational identification. *International Journal of Management, Public Policy and Research, 3*(3), 43-52. <https://doi.org/10.55829/ijmpr.v3i3.229>

Li, B. D., Wang, Y. B., Tang, N. S., & Chen, X. D. (2008). Testing the three-factor model of organizational identification's psychological structure. *Statistics & Decision, 24*(12), 167–169. <https://www.tjyc.com/>

Li, D. (2024). Universities' organizational environment and teachers' organizational citizenship behavior: Evidence from Chinese universities. *Pakistan Journal of Life and Social Sciences, 22*(2), 5263–5276. https://pjlss.edu.pk/pdf_files/2024_2/5263-5276.pdf

Li, F., Jiao, R., Liu, D., & Liu, L. (2023). Principals' transformational leadership and kindergarten teachers' career calling: The mediating roles of organizational identification and leader-member exchange. *SAGE Open, 13*(4), 21582440231218842. <https://doi.org/10.1177/21582440231218842>

Mohamad, M. H., & Majid, A. H. A. (2025). Talent retention among Malaysian private higher education institution academicians through green human resource management. *International Journal of Academic Research in Business and Social Sciences, 15*(6), 618–630. <https://doi.org/10.6007/IJARBS/v15-i6/25728>

Nair, S. G., Sahu, C., & Sadom, N. Z. M. (2024). Mitigating challenges in the administration of private higher educational institutions in Malaysia. *Journal of Quality in Education, 14*(23), 29–43. <https://doi.org/10.37870/joqie.v14i23.420>

Peng, Z. (2025). Research on the current status and challenges of internationalisation in Guangdong's private higher education institutions. *Journal of Sociology and Education, 1*(7). <https://doi.org/10.63887/jse.2025.1.7.16>

Plaku, A. K., Pishkur, E., & Beliu, F. (2025). The role of leaders in shaping school culture: Implications for teachers' motivation and innovation. *Frontiers in Education, 10*, 1541525. <https://doi.org/10.3389/educ.2025.1541525>

Pugliese, E., Bonaiuto, M., Livi, S., Theodorou, A., & van Knippenberg, D. (2024). Team identification and organizational citizenship/counterproductive behavior. *Journal of Applied Social Psychology, 54*(1), 116–125. <https://doi.org/10.1111/jasp.13017>

Rovetta, A., Bortolotti, A., & Palumbo, R. (2025). Integrating team and organizational identity: a systematic literature analysis. *Frontiers in Organizational Psychology, 2*, 1-15. <https://doi.org/10.3389/forgp.2024.1439269>

Sidorenkov, A. V., & Muchinsky, P. (2023). Associations of employees' identification and citizenship behavior in organizations: A systematic review and meta-analysis. *Management*

Review, 73(2), 165–194. <https://doi.org/10.1007/s11301-022-00258-9>

State Council of the People's Republic of China. (2021, May 14). *Implementing Regulations for the Law of the People's Republic of China on the Promotion of Private Education (Decree No. 741)*. https://www.gov.cn/zhengce/content/2021-05/14/content_5606463.htm

Wang, Y. (2025). Perceived organizational fairness, trust, organizational identification, and university teachers' professional development motivation. *PLOS ONE*, 20(1), e0317445. <https://doi.org/10.1371/journal.pone.0317445>

Wang, Y. B. (2005). Organizational identification of employees in state-owned enterprises during transition: Characteristics and causes based on survey data. *New Horizons from Tianfu*, 21(2), 77–84. <http://www.tianfuxinlun.com/ch/index.aspx>

Weisman, H. M., Knippenberg, D., & Sluss, D. (2023). Antecedents of organizational identification: A review and agenda for future research. *Journal of Management*, 49(7), 1915–1954. <https://doi.org/10.1177/01492063221140049>

Wiyono, D., Dewi, D. A., Ambiapuri, E., Parwitasari, N. A., & Hambali, D. S. (2025). Strategic ESG-driven human resource practices: Transforming employee management for sustainable organizational growth in private higher education institutions. *Journal of Higher Education Policy and Management*, 21(1), 65-82. <https://doi.org/10.48550/arXiv.2505.08201>

Wu, J. (2024). Relationship between teachers' professional identity and career satisfaction among college teachers: The role of career calling. *Frontiers in Psychology*, 15, 1-8. <https://doi.org/10.3389/fpsyg.2024.1348217>

Xu, Z., & Pang, N. S. K. (2024). Promoting Teachers' Organizational Commitment: The Effects of Authentic Leadership, Teachers' Well-Being and Social–Emotional Competence. *Behavioral Sciences*, 14(10), 862. <https://doi.org/10.3390/bs14100862>

Yang, J., Nguyen, H.-H. D., Xiong, X., & Wang, X. (2019). Feeling identified vs. behaving as such: A multi-study project on Chinese organizational identification and Chinese employees' identification profiles. *Frontiers in Psychology*, 10, 1039. <https://doi.org/10.3389/fpsyg.2019.01039>

Yang, X., Pu, L., & Shen, S. (2025). Task performance, turnover intention, and their mechanisms in Chinese rural teachers: The mediating role of organizational identity and job satisfaction. *Social Behavior and Personality*, 53(3), e14274. <https://doi.org/10.2224/sbp.14274>

Yang, Y., & Ling, Q. (2023). The influence of existence–relatedness–growth need satisfaction and job burnout of young university teachers: The mediating role of job satisfaction. *Frontiers in Psychology*, 14, 1-9. <https://doi.org/10.3389/fpsyg.2023.1205742>

Yang, Z. M. (2013). *Organizational identification among university teachers and its influencing factors*. [Master's thesis, Huaqiao University]. China National Knowledge Infrastructure. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201501&filename=1014079147.nh>

Zhan, Q., Wang, X., & Song, H. (2023). The relationship between principals' instructional leadership and teacher retention in the undeveloped regions of central and western China: the chain-mediating role of role stress and affective commitment. *Behavioral Sciences*, 13(9), 738. <https://doi.org/10.3390/bs13090738>

Zhou, X. W. (2018). *A study on the status quo and influencing factors of organizational identification among university teachers*. [Master's thesis, Heilongjiang University]. China National Knowledge Infrastructure. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201802&filename=1018015229.nh>

From Learning to Practice: Evaluating Collegial Leadership Training and Transfer Among Malaysian University Academics

ABSTRACT

Authors: Nurul Afiqah Zulkifly^{1*}, Ismi Arif Ismail¹ and 'Siti Noormi Alias

Affiliation: Department of Professional and Continuing Education, Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:
nurulafiqah@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This study examines the effectiveness of a collegial leadership training programme (CLTP) delivered to academics at a Malaysian public university. Using data from one experimental cohort, changes in collegial leadership competencies were assessed using pre- and post-training measures. Training transfer readiness was captured using a validated post-training transfer scale. Repeated measures ANCOVA was conducted to determine the effect of the CLTP on post-training collegial leadership while controlling for demographic factors. Descriptive analyses were used to examine the extent of training transfer readiness. Results indicate significant improvement in collegial leadership competencies, with participants demonstrating moderate-to-high levels of readiness. Regression analysis shows that post-training collegial leadership significantly influences training transfer readiness. These findings highlight the value of structured collegial leadership training for academic staff and underscore the need for institutional support to sustain behavioural change. Implications for training design and leadership development in higher education are discussed.

Keywords: Training Transfer Readiness, Collegial Leadership, Pre- and Post-test, Academic Leaders, University, Malaysia, Kirkpatrick Evaluation Model.

1. Introduction

Collegial leadership has emerged as a pivotal competency among academics within higher education, particularly as institutions navigate increasingly complex and collaborative environments. The shift towards shared decision making, mutual respect, and professional relationships shapes both academic performance and the organisational culture of universities. A clear understanding of collegial behaviours, supported by formal leadership training, can strengthen institutional capacity to address contemporary educational challenges. This is supported by Amtu et al. (2021), which emphasises the role of leadership in fostering collaborative academic cultures that enhance educational quality in higher education institutions. Evidence further suggests that supportive environments characterised by collaboration and shared responsibilities foster professional development and enhance academic performance (Tapic & Baguio, 2024), while collegial leadership enables productive interactions that improve institutional outcomes (Zulkifly, 2023). Such environments are also associated with higher job satisfaction and faculty retention (Singh & Patel, 2025). Nonetheless, there remains a notable gap in formal training focused on collegial leadership skills among academic staff, as leadership development often prioritises managerial competencies over relational and collaborative capabilities (Anumolu & Jain, 2025).

Structured collegial leadership training programmes are increasingly recognised as important mechanisms for developing interpersonal, communication, and collaborative problem-solving skills among academics in higher education. These programmes aim to facilitate shared decision making, foster mutual respect, and strengthen professional relationships within academic communities. Their effectiveness is commonly evaluated using pre-post assessment designs to examine knowledge gains and the application of competencies in practice. Established evaluation frameworks such as Kirkpatrick's model allow for assessment across multiple levels including learning and behavioural change. Prior studies indicate that well designed training interventions can enhance both skill development and self-efficacy in applying learned competencies (Young et al., 2022), while alignment with institutional needs strengthens their relevance and impact (Zulkifly, 2023).

Training effectiveness is closely linked to the extent to which learned competencies are translated into practice. Within higher education, the application of collegial leadership behaviours is shaped by contextual factors such as workload, peer dynamics, and institutional support. Evidence from different contexts shows that pre-post training approaches can lead to improvements in leadership related self-efficacy and collaborative capabilities (Young et al., 2022). Similar patterns have been observed in peer leadership development and mentoring interventions (Merwe et al., 2025), as well as in nursing education where structured training enhances behavioural competencies required for effective leadership (Kelley et al., 2023). However, challenges remain in sustaining these outcomes, as factors such as participant engagement, varying baseline competencies, and contextual constraints may influence the extent to which learning is enacted in practice (Beiting et al., 2023; Kelley et al., 2023). These considerations highlight the importance of examining not only learning outcomes but also participants' readiness to apply collegial leadership behaviours within their professional roles. In addition, behavioural enactment may begin within the training context itself, as participants engage in structured case-based activities, providing early indications of transfer readiness prior to workplace implementation.

Therefore, this study examines the effectiveness of a Collegial Leadership Training Programme (CLTP) conducted at a Malaysian public university. Specifically, it investigates differences between pre-training and post-training collegial leadership, and examines the influence of post-training collegial leadership on training transfer readiness. The programme was designed to enhance academics' engagement in collegial practices such as constructive dialogue, conflict resolution, shared reflection, and collaborative leadership. Guided by Kirkpatrick's Training Evaluation Model, this study focuses on Level 3 behavioural outcomes by examining changes in collegial leadership competencies and their association with perceived transfer readiness. To achieve this, a one group pre-post quasi experimental design is employed to assess within participant change over time. While this design does not permit

strong causal inference due to the absence of a comparison group and potential influences such as maturation or repeated measurement effects, it is appropriate for evaluating training related change in applied academic settings. By linking post-training competencies to perceived transfer readiness, this study provides evidence of early-stage behavioural development consistent with Level 3 of Kirkpatrick's model and addresses the gap in structured collegial leadership training within the Malaysian higher education context.

2. Literature Review

2.1. Collegial Leadership in Higher Education

Collegial leadership represents a relationship-oriented approach that is vital for fostering shared decision-making, mutual respect, and academic autonomy within higher education. This collaborative framework is particularly crucial in public universities, where the practice of distributed leadership can significantly enhance teamwork, curriculum development, and research collaboration. Research has shown that collegial leadership leads to improved communication, effective conflict management, and an enriched culture of academic professionalism, all of which are essential components of a thriving academic environment (Farzana et al., 2022; Jamali et al., 2022).

The alignment of collegial leadership with effective teamwork and academic success emphasizes the necessity for structured training initiatives aimed at enhancing the interpersonal skills and collaborative abilities of academics. However, a significant gap exists in the availability of systematic training specifically focused on collegial practices. As highlighted by Almessabi (2021), there is a pressing need for induction programs tailored to equip faculty with the competencies necessary for practicing collegial leadership effectively. Absent this training, many academics experience inconsistent behaviors and are ill-prepared for leadership roles that demand collaboration and open communication.

Furthermore, studies indicate that the benefits of collegial leadership extend beyond interpersonal relationships, significantly influencing institutional outcomes. For instance, research highlights that distributed leadership models, where authority is shared among faculty, can lead to greater innovation, improved organizational effectiveness, and heightened engagement from staff (Phillips et al., 2024; Jamali et al., 2022). These findings underscore the critical importance of fostering a collegial culture, particularly in environments faced with increased challenges and pressures over recent years (Farzana et al., 2022).

Evaluating the effectiveness of leadership training programs is essential to ascertain whether these interventions appropriately enhance the competencies needed for collegial leadership. Pre-post assessment designs offer a well-established method for evaluating such training, enabling institutions to measure learning gains and the degree of behavioral change in participants' professional settings. Nonetheless, the mere existence of training programs is insufficient; they must be rigorously assessed and tailored to develop the specific relational and collaborative skills vital for academic leadership. Without such evaluations and adjustments, the potential impact of training may remain unrealized, with participants lingering in a state of limited readiness for effective leadership roles (Namutebi, 2024).

In summary, the critical need for structured collegial leadership training in higher education is evident, and an integrated approach to training evaluation is paramount for optimizing the development of essential competencies. As institutions prioritize fostering a collegial culture, they must invest in ongoing training and assessment to equip academics with the tools necessary to thrive in collaborative environments.

2.2. Leadership Development for Academics

Leadership development programs for academics, while essential, often place disproportionate emphasis on administrative and managerial skills at the expense of relational and interpersonal

competencies. This oversight can critically impede the development of leadership that is effective in collaborative and collegial academic environments. Structured leadership programs can significantly enhance various competencies among academics, such as self-awareness, interpersonal communication, decision-making skills, and overall confidence in leadership roles (Zulkifly, 2023). This forms a foundation for more effective collegiality and teamwork in academic settings.

However, despite the demonstrated importance of relational skills, many universities offer leadership development that is informal or experience-based, failing to provide a systematic trajectory for cultivating collegial behaviors (Wennerberg & McGrath, 2022). As Wennerberg and McGrath highlight, structured faculty development interactions can yield substantial personal learning; yet such reflection often remains elusive in traditional teacher training formats (Wennerberg & McGrath, 2022). This gap suggests a critical need for leadership training programs specifically tailored to the academic context, emphasizing skills pertinent to collaboration, professional dialogue, and collegiality.

Effective leadership in academic institutions is intimately connected to how faculty interact with one another. The collegial approach fosters an environment where communication and cooperative problem-solving are prioritized, ultimately enhancing institutional effectiveness (Esterhazy et al., 2021). Research has indicated that successful collaboration can lead to significant organizational improvements, which is evident in the positive impact shown through structured leadership training (Abbas et al., 2024). For example, Abbas et al. (2024) emphasize the transformative effects of in-service training, which fosters not only individual leader effectiveness but also contributes to a more collaborative school climate and stronger collective morale among staff.

A structured approach to leadership training can also address the disparities in readiness for leadership roles among academics. Recent systematic evidence further reinforces this perspective by identifying a set of core leadership competencies essential in higher education contexts, including communication skills, interpersonal abilities, critical thinking and decision-making, emotional intelligence, and digital or ICT-related competencies. These competencies are consistently associated with effective leadership development and reflect the multidimensional nature of academic leadership in contemporary university environments (Saidi & Abd Aziz, 2025). Many faculty members enter their positions with varying levels of experience and preparedness for leadership tasks, which can lead to inconsistent behavior in collaborative settings (Winebrake et al., 2024). By implementing comprehensive training programs that encompass integral relational competencies, institutions can not only improve individual leadership effectiveness but also enhance the overall academic atmosphere.

While administrative competencies are crucial, prioritizing relational and interpersonal skills in leadership development is equally important for fostering effective collegial environments in universities. Structured and contextually relevant training programs are essential to equip academics with the necessary tools for collaboration, enhancing their capabilities for professional dialogue and collegial behavior.

2.3. Training Evaluation and Pre-Post Assessment

Training evaluation is a critical component in determining whether educational interventions successfully achieve their intended outcomes. One of the most recognized frameworks for training evaluation is Kirkpatrick's Training Evaluation Model (1994), which emphasizes four levels: reaction, learning, behavior, and results. This model serves as a foundational guide for assessing the effectiveness of various training programs, especially those focusing on leadership and soft skills, where the core outcomes often hinge upon perceptions and behavioral intentions.

The alignment of evaluation practices with the Kirkpatrick model is evident in various studies, particularly those focusing on academic and leadership training for professionals. For instance,

Douphrate et al. (2024) employed the Kirkpatrick model to demonstrate how Level 3 evaluates trainee behavior change by measuring the application of new skills on the job. This level is crucial as it highlights not only whether the training was effective in a theoretical sense but also its practical implications in participants' work environments. Additionally, Cook et al. (2024) utilized a similar four-level approach to assess the leadership development of medical residents, illustrating the universality of this model across educational domains.

Pre-post assessment designs are widely utilized within this framework to objectively measure improvements in competencies. These designs are particularly well-suited for leadership training, as they allow educators to capture data on participants' perceptions before and after the training interventions, quantifying learning gains effectively (Hamid et al., 2021). Such assessments reveal how well participants absorb the training content and subsequently transfer new knowledge and skills into their daily professional practices. However, some studies highlight that many evaluations do not adequately measure Level 3 (Behavior), indicating a gap in assessment practices (Schaller et al., 2023).

Moreover, the importance of evaluating higher levels of the Kirkpatrick model, specifically Level 3 (Behavior), cannot be overstated. Historical challenges such as time constraints and perceived complexities often inhibit the thorough evaluation of these levels in training programs (Douphrate et al., 2024). Yet, successful outcomes at this level lend credibility to the training initiatives, demonstrating tangible benefits not only for participants but also for the organizations they serve. For example, research by Farhadi et al. (2023) effectively traced competency development in neonatal resuscitation training over time, showcasing the importance of sustained behavioral change as a critical measure of training success.

Utilizing Kirkpatrick's Training Evaluation Model for assessing leadership development programs facilitates a comprehensive understanding of training effectiveness. By focusing not only on reactions and learning but also on behavioral changes and organizational results, institutions can ensure their training initiatives truly prepare academics for the challenges of collaborative leadership, ultimately fostering a more collegial and effective academic environment.

2.4. Training Transfer in Academic Contexts

Training transfer, defined as the application of newly acquired knowledge and behaviors in the workplace, is a critical factor in evaluating the effectiveness of educational interventions, particularly in higher education. Transfer is inherently complex in academic settings due to the influence of autonomy, organizational culture, workloads, and peer dynamics. Research has established that factors such as trainee motivation, work-environment support, and perceived training relevance significantly affect the transfer process, aligning with Baldwin and Ford's framework Baldwin & Ford (1988). Understanding whether academics can effectively apply collegial leadership competencies after training is vital for assessing the long-term value of such interventions.

Studies have emphasized the impact of various contextual factors on training transfer. Cowman & McCarthy (2016) noted that while training transfer is crucial for organizational effectiveness, there remains a limited understanding of how demographic and situational factors influence this transfer. Their findings suggest that these factors can create variability in transfer efficacy, potentially affecting academic initiatives. Consequently, universities must consider how these variables interact within their unique organizational climates.

Motivation to transfer learning back to the job is primarily contingent on trainees' expectations regarding the utility of training. In addition, engagement in structured experiential activities has been shown to enhance participants' readiness to apply learned competencies. Empirical evidence suggests that participation in organized programmes such as volunteering or practice-based initiatives contributes to increased readiness, perceived relevance, and active involvement, all of which are critical precursors to effective training transfer. This highlights the importance of integrating experiential

elements within training design to strengthen the likelihood of behavioural application (Saidi et al., 2025). Kontoghiorghes (2008) proposed a systemic model that suggests employees are more driven to learn and apply new skills if they perceive that their training will positively impact their job performance. This aligns with findings from Yaghi and Bates (2020), which demonstrated a positive correlation between peer support and training transfer in higher education institutions. The connection between robust support systems and the successful application of training underscores the necessity for higher education institutions to foster an environment conducive to transfer.

Furthermore, Cromwell & Kolb (2004) assert that organizational factors, such as promotional opportunities and supportive leadership, play significant roles in facilitating or hindering transfer. Their assertion emphasizes that beyond immediate training programs, institutional support mechanisms are crucial for ensuring that learning is translated into practice. Similarly, Saks & Burke (2012) concluded that transfer of training involves not only the initial learning period but also the maintenance and generalization of these new skills over time.

The interplay between workload and training transfer has also garnered attention. Hughes et al. (2016) highlighted that excessive workload can pose challenges to the application of new skills, suggesting that organizations need to actively manage workload and provide opportunities for application to enhance transfer. However, their meta-analysis suggests that the acuity of a unit's patients may not significantly impact healthcare team training effectiveness, indicating that workload impacts might vary based on context. Grossman & Salas (2011) further delineated the variables related to trainee characteristics, training design, and work environment as predictors of training transfer success.

Hence, effective training transfer in higher education is influenced by a confluence of factors, including the organizational culture, supervisor and peer support, and the perceived relevance of training content. To enhance the application of collegial leadership competencies post-training, institutions must acknowledge these complexities and invest in supportive environments that facilitate ongoing learning and application of skills in real-world contexts. In the present study, training transfer is operationalised as training transfer readiness, reflecting participants' perceived ability to apply learned behaviours immediately following the training, rather than confirmed workplace enactment over time.

2.5. The Underpinning Theory

This study adopts a practical evaluation approach based on Kirkpatrick's model, focusing on changes in collegial leadership competencies and transfer behaviors. The Kirkpatrick model remains a robust framework for assessing training effectiveness, particularly in educational contexts, as it evaluates levels of reaction, learning, behavior, and results (Sim & Radloff, 2008; Alghanaim et al., 2025). This model enables the examination of how training interventions influence learning outcomes and facilitate behavior transfer, which is crucial for determining the long-term value of academic leadership development initiatives. In this context, incorporating demographic and contextual factors as controls enhances the validity of the interpretation of results. Prior research highlights the significance of contextual variables in influencing the transfer of training, affirming that customization to specific situational factors can lead to more meaningful assessments of training efficacy (Zulkifly, 2023).

The conceptual model in this study positions the training intervention as a catalyst for learning, suggesting that effective training can lead to increased self-efficacy and greater application of leadership skills in real-world settings. Studies show that training perceived as relevant within the specific organizational context tends to increase the likelihood of successful transfer behaviors (Alghanaim et al., 2025).

Furthermore, empirical evidence describes how structured training environments, when effectively designed, can positively impact learners' ability to apply new competencies, reaffirming the importance of transferring learned skills back to their work environments. Training that emphasizes practical, collegial interactions fosters a supportive learning climate that enhances participant engagement and

motivation to transfer skills learned in professional settings (Sim & Radloff, 2008; Maulidya et al., 2024; Alghanaim et al., 2025).

By leveraging Kirkpatrick's model with a focus on contextual and demographic variables, this study aims to provide a comprehensive evaluation of the effectiveness of collegial leadership training programs in higher education. Assessing both learning outcomes and behavior changes will shed light on the efficacy of these programs in fostering lasting improvements in leadership competencies among academics.

2.6. Research Framework

The research framework of the study is presented in Figure 1.

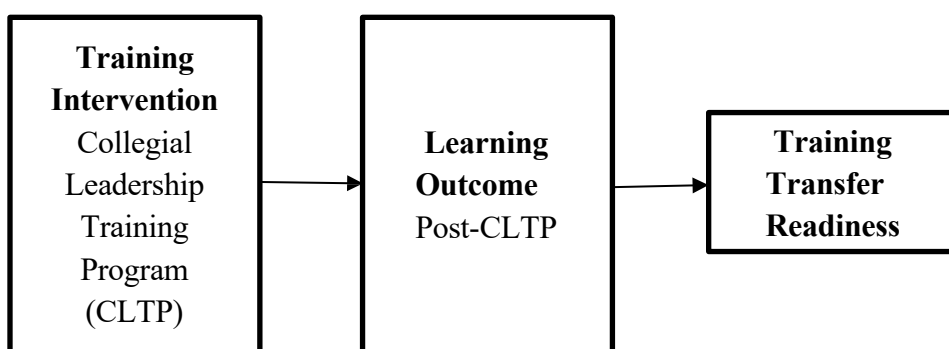


Figure 1. Research framework

2.7. Hypothesis of the Study

Based on the preceding discussion, the following hypotheses are proposed:

H₁: There is a significant difference between pre- and post-training collegial leadership readiness.

H₂: Post-training collegial leadership significantly influences training transfer readiness.

2.8. Research Gap and Study Contribution

Despite the growing body of research on collegial leadership and training transfer in higher education, several empirical gaps remain. First, much of the existing literature examines leadership development outcomes at the level of self-reported learning or satisfaction, with comparatively limited attention to behavioural evidence aligned with Kirkpatrick's Level 3 (Behaviour/Transfer), particularly within academic settings (Wißhak et al., 2025; Sharif et al., 2022). Second, while Baldwin and Ford's transfer framework has been widely cited, empirical studies that explicitly link post-training learning outcomes to subsequent transfer behaviours among university academics remain scarce (Chauhan et al., 2024). This is especially evident in non-Western higher education contexts, where institutional structures, collegial norms, and leadership practices may differ substantially (Nengyanti et al., 2024).

Moreover, prior studies tend to conceptualise training transfer as an outcome assessed solely after training completion, overlooking the role of behavioural practice during training itself (Leránóz-Iglesias et al., 2023). There is limited empirical attention to "in-training transfer," where participants actively enact collegial leadership behaviours through case-based or simulated academic scenarios that closely resemble their professional contexts. As a result, how behavioural practice during training relates to subsequent workplace transfer remains underexplored (Ng et al., 2024).

Addressing these gaps, the present study contributes empirical evidence from a Malaysian public university by integrating pre-post assessments of collegial leadership competencies with post-training transfer measures aligned to Kirkpatrick's Level 3 evaluation. By examining both learning gains and their relationship to training transfer, and by recognising behavioural practice embedded within the training process, this study extends existing leadership and training transfer research within higher education and provides contextually grounded evidence of how collegial leadership can be developed and enacted in academic work environments.

3. Methodology

3.1. Research Design

This study employed a quasi-experimental pre-post one-group design to evaluate the effectiveness of the Collegial Leadership Training Programme (CLTP). This design is particularly appropriate for assessing learning improvements and behavioral transfer following a structured intervention, aligning well with Kirkpatrick's Level 3 (Behavior/Transfer). The central objective of this evaluation was to determine whether participants of one experimental cohort experienced improvements in their competencies over time and whether they applied their training in their respective workplaces.

Kirkpatrick's model serves as a guiding framework in this evaluation process whereby Level 3 examines the degree to which this learning translates into behavioral changes in the workplace. The quasi-experimental design, specifically the pre-post assessment, allows for a clear measurement of the competencies before and after the intervention, thereby providing insights into the effectiveness of the training program.

Utilizing a one-group pre-and post-test quasi-experimental design facilitates the measurement of changes in specific leadership competencies that the training aimed to enhance. Previous studies have demonstrated the utility of this design in educational settings, especially in evaluating similar leadership training programs. The absence of a comparison group does present limitations; however, the focus on individual progress allows for a robust evaluation of how well the training meets its objectives in fostering collegial leadership skills.

Additionally, controlling for demographic factors (gender, age, qualification, administrative position, and duration of service), is crucial for ensuring a comprehensive understanding of the results. By accounting for these variables, the evaluation aims to mitigate confounding effects that may influence learning outcomes and behavioral transfer. This approach not only strengthens the validity of the findings but also enhances our understanding of the broader context in which the training occurs, which is essential for translating training into practical applications in academic settings.

In summary, this study's design and evaluation framework reflect a commitment to understanding the effectiveness of the collegial leadership training program (CLTP) in improving collegial leadership competencies. By employing Kirkpatrick's model, the research aims to demonstrate participants' readiness to transfer learned behaviours into real-world applications, thus contributing meaningful insights into leadership development in higher education.

3.2. Methodological Limitations and Inferential Boundaries

This study adopts a one-group pre-post quasi-experimental design to evaluate the CLTP within an authentic academic setting. While this design is appropriate for applied training evaluation (Creswell & Creswell, 2017), it is subject to several recognised threats to internal validity that necessarily constrain causal interpretation. First, history effects may have influenced participants' post-training responses. Events occurring concurrently with or shortly after the training, such as departmental discussions, institutional initiatives, or leadership related experiences, may have contributed to observed changes independently of the intervention.

Second, testing effects may have arisen from repeated exposure to the collegial leadership instrument, whereby participants' increased familiarity with the items heightened awareness of collegial behaviours rather than reflecting actual behavioural change. Third, maturation effects cannot be ruled out, as participants may have naturally developed greater leadership awareness or confidence over time irrespective of the training. Fourth, the possibility of a Hawthorne effect exists, whereby participants modified their responses or behaviours simply due to their awareness of being involved in a training programme and subsequent evaluation (Creswell & Creswell, 2017).

In light of these threats, the findings are interpreted cautiously and are not presented as definitive causal evidence of training effectiveness. Rather, the results are framed as indicative of training associated change within participants over time. By explicitly acknowledging these inferential boundaries, the study situates its findings appropriately within the training evaluation literature, providing contextually grounded evidence of behavioural development while avoiding overstatement of causal claims.

3.3. Participants

The study involved a group of 30 academics from a Malaysian public university who voluntarily participated in the CLTP. Pre- and post-training assessments were utilized to evaluate improvements in collegial leadership competencies, and various demographic factors were collected to analyze their potential influence on training outcomes. The demographic information gathered from participants included gender, age, qualification, administrative position, and duration of service.

Such demographic considerations are essential in training evaluations, as they can help interpret the results in the context of different academic backgrounds and experiences. For example, demographic factors like age and gender may influence the readiness to adopt new leadership practices or motivation levels to engage actively in training (Cadena-Povea et al., 2025).

Participation in this training program was entirely voluntary, and all participants provided informed consent, which underscores the ethical integrity of the research process. The selection of a voluntary sample may contribute to a higher level of engagement and motivation among participants, potentially resulting in more pronounced training effects (Ladyshevsky & Flavell (2011).

Further, applying Kirkpatrick's model in this context, specifically focusing on Level 3 (Behavior/Transfer), is crucial to ascertain not only whether the participants learned new competencies but also whether they effectively applied these skills within their professional environments after training. Previous research supports the assertion that demographic factors, such as academic rank and years of service, may also moderate the transfer of training, highlighting the importance of controlled analysis to draw meaningful conclusions (Zulkifly et al., 2021).

Ultimately, this research aims to contribute to the understanding of how structured leadership development initiatives can affect academic professionals in Malaysian public universities, paving the way for enhanced collegiality and collaboration in academic settings.

3.4. Training Intervention

The Collegial Leadership Training Programme (CLTP) comprises a structured session aimed at enhancing participants' collegial leadership competencies, focusing on key areas such as communication, collaborative problem-solving, reflective dialogue, and conflict management. The training program is designed with careful documentation of content, duration, facilitators, and learning activities to ensure consistency and facilitate effective evaluation.

The emphasis on communication and collaborative problem-solving through case study aligns with findings from Ayeleke et al. (2019), which noted that delivering training in a context relevant to participants' everyday work conditions significantly enhances the application of learned skills. The presence of situational cues and consequences within training settings increases the likelihood that participants will transfer skills back to their workplace. This principle is particularly vital in higher education, where fostering positive peer dynamics and effective communication can lead to improved organizational culture and performance.

Additionally, evaluation of training effectiveness is essential. Tan (2024) highlighted the importance of applying evaluation models to assess the impacts of training, considering potential weaknesses within specific cultural contexts. Such evaluation frameworks are critical for understanding how well training is received and its subsequent transfer into practice among academics. Therefore, the CLTP aims not only at knowledge acquisition but also at the sustainable application of these competencies in collaborative academic environments.

Although Kirkpatrick's Level 3 evaluation traditionally examines behavioral change in the workplace after training, contemporary training evaluation literature recognizes that behavioral application can also occur during the training when participants engage in realistic practice activities. In this program, participants completed structured case-based exercises that required them to apply collegial leadership principles to authentic academic scenarios. Through discussion, role interpretation, collaborative decision-making, and problem-solving, participants demonstrated behavioral practice of the competencies introduced in the training. These activities constitute a form of "in-training transfer," where learners apply skills in simulated environments that closely resemble their actual work settings. This idea is supported by Douphrate et al. (2024), who state that Level 3 measures behavior change by assessing how trainees apply their newly acquired knowledge and skills on the job.

Therefore, the case study discussions provide valid behavioral evidence aligned with Level 3 of Kirkpatrick's model. The significance of in-training transfer is also discussed by Walter et al., who indicate that experiential learning and simulations within training programs effectively enhance teamwork skills, demonstrating the relevance of realistic practice during training (Walter et al., 2024). Additionally, Boet et al. note that learning occurs in simulated environments, affirming that such approaches effectively mirror the conditions of actual work scenarios (Boet et al., 2014).

The structured exercises within the CLTP enable participants to engage proactively with the content, facilitating a deeper understanding and enabling the transfer of skills as advocated by the principles of Kirkpatrick's evaluation framework. This approach resonates with the findings of Miller et al. (2012), who report improvements in teamwork and communication in clinical settings through simulation training. Overall, the emphasis on behavioral practice within training contexts reflects a modern understanding of training evaluation that recognizes the dynamic interplay between learning and behavioral application during training sessions.

Furthermore, interprofessional learning is an increasingly acknowledged necessity in leadership development, as underscored by McGowan et al. (2020). This is particularly important in educational settings due to the collaborative nature of academic work, which necessitates strong communication and teamwork skills. The training incorporated in the CLTP nurtures these essential interpersonal skills while ensuring flexibility to adapt to evolving educational needs.

In conclusion, the strategic design and implementation of the CLTP prioritize enhancing collegial leadership competencies while also recognizing the significance of contextual relevance and structured evaluation. By fostering an environment conducive to peer support and effective communication, the training program aims to achieve meaningful and sustainable improvements in academic leadership effectiveness.

3.5. Instruments

(a) Collegial Leadership (Pre-Post)

Collegial leadership was measured using a five-point Likert scale instrument developed by Seigel and Miner-Rubino (2009), namely the Collegiality Climate Scale (CCS). The sample item for this instrument is “I am willing to help each other.” The Collegiality Climate Scale (CCS) was originally developed within a Western higher education context, specifically in a United States law school setting. Although developed in a different context, the CCS captures general collegial behaviours such as collaboration, communication, and mutual support, which are applicable across academic environments. For this study, items were reviewed for contextual relevance, and minor wording adjustments were made to ensure suitability for the Malaysian university setting.

(b) Training Transfer Readiness

Training transfer was measured using a five-point Likert scale instrument developed by Facticeau et al. (1995). A sample item of this instrument is “I am able to transfer the skills learned from training courses back to my actual job.” In this study, training transfer is operationalised as participants’ perceived readiness to apply learned collegial leadership behaviours immediately following the training, rather than confirmed workplace enactment over time. Minor wording adaptations were made to ensure contextual relevance to the Malaysian university setting, including replacing references to “state government” with “university.”

(c) Covariates

This study controls for demographic factors which include gender, age, qualification, administrative position, and duration of service.

3.6. Ethical Consideration

Ethical approval for this study was obtained from the Jawatankuasa Etika Universiti Putra Malaysia (JKEUPM), Universiti Putra Malaysia (Reference No. JKEUPM-2019-356). Participation in the study was voluntary, and informed consent was obtained from all participants prior to data collection. Participant email addresses were collected solely for the purpose of matching pre-training and post-training responses. These identifiers were removed after matching and were not included in the final dataset used for analysis. The data were subsequently de-identified and treated as confidential. All research data were stored securely on password protected devices and were accessible only to the research team. Findings are reported in aggregate form to minimise the risk of participant identification.

3.7. Data Collection Procedure

The data collection for this study was executed in three distinct phases, allowing for a comprehensive evaluation of the Collegial Leadership Training Programme (CLTP). The phases included:

1. **Pre-training Assessment of Collegial Leadership:** This initial phase involved gathering baseline data on participants' collegial leadership competencies. An online questionnaire, administered via a QR code, provided a clear starting point regarding participants' skills and self-perceptions before the training intervention.
2. **Training Delivery (CLTP):** Following the pre-training assessment, the CLTP was delivered with structured content designed to enhance critical collegial leadership competencies,

including communication, collaborative problem-solving, reflective dialogue, and conflict management. This structured approach facilitates deep engagement with relevant material, aligning with findings from Arabi and Garza (2022), which emphasized that well-designed training contributes to effective transfer of training to the workplace.

3. **Post-training Assessment and Training Transfer Readiness Questionnaire:** The final phase involved reassessing participants' competencies and gathering data on the training transfer readiness in relation to their professional environments through an online questionnaire. This post-training assessment allows for direct measurement of learning gains, while the training transfer readiness questionnaire evaluates participants' application of new skills in practice. This acknowledges the importance of contextual factors in fostering effective transfer, as identified by Martin (2010). Both the post-training assessment and the training transfer readiness questionnaire were administered immediately after completion of the training programme to capture initial behavioural enactment and readiness for transfer consistent with Level 3 (Behaviour) evaluation, with case-based activities used during the training to provide participants with opportunities to practise and demonstrate collegial leadership behaviours in realistic academic scenarios.

All questionnaires were administered online via a QR code, ensuring an efficient and accessible process for participants. The online format aligns with modern training evaluation environments, facilitating quick feedback and data collection processes. In conclusion, the structured evaluation design, based on pre- and post-assessments combined with a training transfer readiness questionnaire, provides insight into how effectively the CLTP influences participants' leadership competencies and their ability to transfer these skills into their academic work. This approach aids in understanding the immediate impacts of training and provides essential data on the long-term effectiveness of leadership development initiatives in higher education.

3.8. Data Analysis

The data analysis for this study was conducted using SPSS, providing a robust statistical foundation for evaluating the effectiveness of the Collegial Leadership Training Programme (CLTP). The analysis included several key components:

- **Reliability Assessment:** Cronbach's alpha values were computed to assess the internal consistency of the measurement scales used in the questionnaires. This step is crucial for ensuring the reliability of the data being analyzed, reflecting the extent to which the items within each scale yield consistent results across respondents. Prior research highlights the importance of establishing reliability in training assessments, as it directly relates to the legitimacy of the findings (Fernandez et al., 2016; Lacerenza et al., 2017).
- **Repeated measures ANCOVA:** The primary statistical method employed was ANCOVA, where the post-training collegial leadership competencies served as the dependent variable (DV). The demographic factors were treated as controls. Utilizing ANCOVA is a well-supported method in training research, enabling researchers to account for pre-existing variations and to ascertain the training's actual effect on the changes in competencies post-intervention (Ng, 2014).
- **Training Transfer Readiness Evaluation:** Descriptive statistics, including means, standard deviations, and distributions, were computed to assess the training transfer readiness reported by participants. This analysis serves to quantify the degree to which participants felt they could apply the skills learned during the training in their workplace settings. Understanding the

nature of training transfer is essential, as it can reveal whether the training ultimately contributes to effective changes in academic leadership behaviors, aligning with findings from studies on training transfer in similar contexts (Ng, 2014; Zulkifly et al., 2021).

- **Hierarchical Regression Analysis:** A hierarchical regression analysis was conducted to explore the influence of post-training collegial leadership on training transfer readiness. This method provides insight into how post-training collegial leadership relates to training transfer readiness. Such analyses can illuminate the connections between training outcomes and real-world applications, supporting the broader understanding of training effectiveness (Sparr et al., 2016).

An a priori power calculation was not conducted because the study evaluated one intact training cohort within an operational university setting. To support the adequacy of the sample for the primary analyses, post hoc consideration of the observed effect sizes was undertaken. The repeated measures analysis demonstrated a large training related effect, and the regression model predicting training transfer readiness explained a substantial proportion of variance. These results indicate that the study was sufficiently sensitive to detect moderate to large effects. However, the sample size may have limited sensitivity to detect small effects or subgroup differences, and nonsignificant findings should therefore be interpreted cautiously. Quasi experimental studies require a minimum of 15 participants in each group (Creswell & Creswell, 2017) for the purpose of statistical procedures, Tabachnick and Fidell (2007) suggested that each cohort consists of a minimum of 20 participants to ensure robustness in the results of the paired sample t-test and analysis of covariance (ANCOVA).

Assumptions for the repeated measures ANCOVA were assessed. Residual diagnostics indicated approximate normality. Homogeneity of regression slopes was supported by nonsignificant Time by covariate interactions. As the within participant factor comprised two levels, sphericity was not applicable. Regression diagnostics indicated no violations of key assumptions. Standardized residuals were within acceptable ranges, and inspection of Cook's distance, leverage values, and Mahalanobis distance did not identify influential cases or multivariate outliers.

In summary, the data analysis process employed a comprehensive approach, utilizing reliability testing, ANCOVA, descriptive statistics, and optional regression analysis to assess the effects of the CLTP on collegial leadership competencies and training transfer readiness among the participants. This multi-faceted analytical strategy ensures that the study's conclusions are grounded in robust statistical evidence.

4. Findings and Discussion

4.1. Reliability Analysis

Table 1 presents the reliability coefficients for all measurement scales. The collegial leadership scale demonstrated high internal consistency at both pre-training ($\alpha = .903$) and post-training ($\alpha = .898$). The training transfer readiness scale also showed satisfactory reliability ($\alpha = .925$).

Table 1. Reliability Coefficients for Study Variables

Variable	Number of Items	Cronbach's α
Collegial Leadership Training Program (Pre)	11	.903
Collegial Leadership Training Program (Post)	11	.898
Training Transfer Readiness	9	.925

4.2. Descriptive Statistics

Table 2 shows the descriptive statistics for the pre- and post-training collegial leadership scores. Mean post-training scores were higher than pre-training scores, indicating improvements following the intervention.

Table 2. Descriptive Statistics for Collegial Leadership (Pre & Post)

Variable	Mean	SD	Min	Max
Pre-training CL	4.245	.494	2.80	5
Post-training CL	4.542	.399	3.64	5
Training Transfer Readiness	4.196	.509	3.11	5

4.3. Repeated Measures ANCOVA for Training Effectiveness

A repeated measures ANCOVA was conducted to examine whether the training had a significant effect on post-training collegial leadership after controlling for demographic factors (gender, age, qualification, administrative position, and duration of service). The results are discussed below.

4.3.1. Pairwise Comparison

Pairwise comparisons using Bonferroni adjustment showed a significant increase in collegial leadership scores from pre-test to post-test: Mean difference = -0.297 ; SE = 0.079 ; $p < .001$. This confirms that the training yielded statistically significant improvement. Table 3 shows pairwise comparison results.

Table 3. Pairwise Comparisons Between Pre-test and Post-test Scores

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% CI Lower	95% CI Upper
1 (Pre)	2 (Post)	-.297*	.079	<.001	-.459	-.135
2 (Post)	1 (Pre)	.297*	.079	<.001	.135	.459

* Significant at .05 level

4.3.2. Multivariate Test

The multivariate repeated-measures tests indicate a statistically significant difference across time, confirming that the training improved collegial leadership learning even after adjusting for all covariates (Pillai's Trace = $.373$; $F(1,24) = 14.307$; $p < .001$; Partial $\eta^2 = .373$ - large effect). This large effect size indicates that the training intervention was highly impactful.

Table 4. Multivariate Tests for the Effect of Time

Test	Value	F	Hypothesis df	Error df	Sig.	Partial η^2
Pillai's Trace	.373	14.307	1	24	<.001	.373
Wilks' Lambda	.627	14.307	1	24	<.001	.373
Hotelling's Trace	.596	14.307	1	24	<.001	.373
Roy's Largest Root	.596	14.307	1	24	<.001	.373

4.3.3. Tests of Within-Subjects Effects (Interactions with Covariates)

Interaction terms (Time $\times$ Covariates) test whether improvement over time differed across demographic groups. According to the results, none of the interactions are significant, all p -values $> .157$ and partial η^2 values are very small ($< .081$). This means the training improvement was consistent across gender, age, qualification, administrative position, and duration of service.

Table 5. Tests of Within-Subjects Effects

Source (Interaction)	F	df	Sig.	Partial η^2
Time $\times$ Gender	1.534	1	.227	.060
Time $\times$ Age	.298	1	.590	.012
Time $\times$ Qualification	.006	1	.938	.000
Time $\times$ AdminPos	.605	1	.444	.025
Time $\times$ Duration of Service	2.129	1	.157	.081

4.3.4. Tests of Between-Subjects Effects (Main Effects of Covariates)

Between-subject's tests show whether demographic covariates influenced overall levels of learning. Based on the results, none of the covariates significantly predicted learning, all p -values are $> .282$, and All effect sizes are extremely small ($\eta^2 < .048$). Therefore, the training was equally effective across demographics. However, these non-significant findings (Table 5 and Table 6) should be interpreted with caution, as the relatively small sample size may limit the statistical power to detect subgroup differences, and therefore do not necessarily indicate equivalence across groups.

Table 6. Tests of Between-Subjects Effects

Covariate	F	df	Sig.	Partial η^2
Gender	1.212	1	.282	.048
Age	.422	1	.522	.017
Qualification	.495	1	.488	.020
Administrative Position	.801	1	.380	.032
Duration of Service	1.184	1	.287	.047

4.4. Hierarchical Regression Predicting Training Transfer Readiness

A hierarchical regression analysis was conducted to examine whether post-training collegial leadership predicted training transfer readiness beyond baseline levels. Baseline collegial leadership was entered in the first step and accounted for 16.1% of the variance in training transfer readiness, $R^2 = .161$, $p = .028$. In the second step, post-training collegial leadership was added to the model, resulting in a significant increase in explained variance, $\Delta R^2 = .179$, $p = .012$. In the final model, post-training collegial leadership emerged as a significant predictor of training transfer readiness ($\beta = .523$, $p = .012$, 95% CI [.161, 1.171]), whereas baseline collegial leadership was no longer significant. Collinearity diagnostics indicated no multicollinearity concerns. These findings provide Level 3 evidence of training effectiveness, consistent with Baldwin and Ford's transfer process model, demonstrating that learning outcomes directly influence workplace application.

Table 7. Hierarchical Regression Predicting Training Transfer Readiness

Predictor	B	SE B	95% CI for B	β	t	p	VIF
Constant	0.762	0.928	-1.141, 2.666		0.822	.418	
Pre-CLTP	0.097	0.199	-0.312, 0.505	.094	0.485	.631	1.527
Post-CLTP	0.666	0.246	0.161, 1.171	.523	2.704	.012	1.527

Model statistics:

- Model 1: $R^2 = .161$, $F(1, 28) = 5.361$, $p = .028$
- Model 2: $R^2 = .340$, $F(2, 27) = 6.951$, $p = .004$
- $\Delta R^2 = .179$, $\Delta F(1, 27) = 7.310$, $p = .012$

Note. Dependent variable = Training transfer readiness. Standardized coefficients (β) are reported. Confidence intervals are based on 95% limits. No multicollinearity issues were detected ($VIF < 2$).

4.5. Discussion

While these findings indicate positive changes following the training, they should be interpreted within the boundaries of the study design. The one group pre-post evaluation captures within participant change over time but does not permit strong causal attribution at the institutional level. Accordingly, the observed improvements are interpreted as training associated gains among the participating cohort rather than definitive evidence of organisational wide impact. The repeated measures ANCOVA results indicated a significant increase in post-training collegial leadership scores after controlling for demographic factors. This suggests that the training successfully enhanced participants' knowledge, awareness, and behavioral readiness related to collegial leadership, aligning with studies indicating that targeted leadership development programs can strengthen interpersonal and collaborative competencies among academics. The empirical evidence supports the notion that such programs play a pivotal role in equipping academics with the necessary skills to navigate complex interpersonal dynamics and promote collaboration effectively.

An additional contribution of this study lies in its attention to in-training behavioural enactment through case-based activities. Rather than conceptualising transfer solely as a post-training outcome, as traditionally emphasised in training transfer frameworks (Baldwin & Ford, 1988), the findings suggest that elements of transfer may begin to emerge within the training context itself, as participants engage in simulated collegial scenarios. This extends traditional perspectives by highlighting the role of structured learning environments in facilitating early-stage behavioural application. This perspective complements existing transfer models by suggesting that behavioural application can be initiated prior to workplace implementation. This is further supported by studies demonstrating that experiential engagement through structured programmes enhances participants' readiness and capacity to apply learned skills. Such findings reinforce the argument that behavioural enactment can begin within the training environment itself, particularly when participants are actively involved in practice-oriented and contextually relevant activities. The successful enhancement of these competencies is consistent with findings from previous research indicating that leadership training, particularly when focused on practical application, can lead to significant improvements in communication, teamwork, and overall academic productivity (Hayes & Ingram, 2019). Specifically, the effective engagement with case-based practices allows for a deeper understanding and application of leadership principles in real-world academic settings, reinforcing the idea that experiential learning is essential for internalizing the behaviors required for successful collegial leadership.

The descriptive analysis of training transfer readiness further revealed that participants reported moderate-to-high levels of readiness to apply the learned collegial leadership skills in their academic roles. This indicates that the training supported participants' perceived ability to enact collegial behaviours such as improved communication, collaborative decision-making, and constructive conflict management. This is consistent with early-stage behavioural outcomes associated with Kirkpatrick's Level 3 (Behaviour), which emphasizes real-world application of learned skills. This interpretation is further supported by the regression analysis, which showed that post-training collegial leadership uniquely predicted training transfer readiness even after controlling for baseline levels, suggesting that training transfer readiness was more strongly associated with training related gains than with pre-existing collegial dispositions.

Several alternative explanations should also be considered when interpreting these outcomes. Improvements in post-training scores may partially reflect testing effects, whereby repeated exposure to the measurement instrument increased participants' awareness of collegial leadership behaviours. A Hawthorne effect may also be present, as participants' engagement in a structured leadership programme and evaluation process could temporarily heighten motivation or responsiveness. In addition, the immediate post-training measurement may capture short term novelty effects rather than sustained behavioural change. These factors do not negate the observed improvements but suggest that the findings reflect early-stage behavioural enactment rather than enduring transformation.

The extent to which collegial leadership behaviours are sustained beyond the training context is likely influenced by organisational readiness and practical constraints within academic environments. Factors such as academic workload, time pressures, departmental culture, and the level of leadership support may shape opportunities for applying collegial practices in daily work. Without supportive structures, participants may find it challenging to consistently enact collaborative decision making and reflective dialogue, despite positive learning outcomes. This suggests that training alone may be insufficient unless accompanied by enabling organisational conditions.

A closer examination of item-level patterns indicates variation in the magnitude of change across different aspects of collegial leadership. Items with relatively high pre-training scores showed smaller gains, suggesting potential ceiling effects where participants had already demonstrated these behaviours prior to the intervention. In contrast, items with lower baseline scores exhibited larger improvements, indicating that the training may have been particularly effective in strengthening less developed competencies such as structured dialogue, collaborative problem-solving, and reflective practices.

5. Conclusion

In examining the impact of a collegial leadership training program on participants' competencies and perceived transfer readiness within academic workplaces, it is crucial to analyze the underlying mechanisms whereby structured training influences relational leadership practices. This study employs a pre-post evaluation design based on Kirkpatrick's model, offering evidence on the effectiveness of such training initiatives in enhancing collegial leadership competencies.

The evaluation tactics used in leadership training frameworks can enhance collegiality by creating an environment where relational competencies are emphasized and practiced. The concept of collegiality and collaboration emerges as a dynamic component in academic settings and is essential in fostering an environment conducive to shared leadership, where structured training programs can play a role in shaping the relational dynamics necessary for academic practice.

This study further shows that academic leadership training is associated with participants' readiness to apply relational skills, suggesting that a structured educational framework can support collaborative practices among higher education staff. This relationship between training and leadership practices highlights the importance of institutional support for such programs in facilitating competence development in academic settings. Future longitudinal research is needed to examine whether these competencies are sustained over time.

In a nutshell, the Collegial Leadership Training Program (CLTP) enhances participants' leadership capabilities and supports the development of relational competencies essential in academic settings. Through structured evaluations grounded in established frameworks such as Kirkpatrick's model, perceived transfer readiness can be assessed within the immediate post-training context, reinforcing the significance of relational leadership practices in academia.

6. Implications and Recommendations

6.1. Theoretical Implications

This study provides empirical support for the relevance of collegial leadership as a competency in Malaysian public universities, particularly in enhancing academic collaboration and communication. By establishing collegial leadership as an essential competency, this study contributes to a growing body of literature emphasizing the need for relational skills in educational environments where teamwork and shared decision-making are paramount.

The findings also contribute to the training evaluation literature by demonstrating that behavioral indicators can be captured during training through case-based practice, not only through post-training assessments. Traditional evaluations often emphasize post-training performance, but the ability to measure behavioral changes during training contexts provides new insights into effective teaching methodologies. This approach aligns with findings that support the use of authentic, context-specific case studies to enhance the internalization of leadership behaviors. The study's results underscore the importance of situational engagement in learning, thereby reinforcing models of situated learning and experiential leadership development, which advocate for transformative learning experiences in real or simulated contexts.

6.2. Practical Implications

From a practical perspective, the findings suggest that structured collegial leadership training can support academics' readiness to practise collaborative behaviours, but such training should not be assumed to automatically translate into sustained workplace change. Human Resource and training units may play a facilitative role by aligning leadership development initiatives with organisational conditions, such as workload allocation, peer support mechanisms, and leadership endorsement. However, institutions should avoid assuming that short duration training alone can resolve deeper structural or cultural challenges affecting collegial practice.

Furthermore, Human Resources and Training Units in Malaysian public universities can use pre-post testing as a straightforward method to track competency growth among academics. This method allows institutions to quantitatively assess the effectiveness of their leadership development initiatives, ensuring that training investments yield tangible benefits. The CLTP demonstrates potential for scaling up as a cost-effective, short-duration intervention that can improve relational dynamics in academic departments (Ladyshevsky & Flavell, 2011). By keeping training interventions concise yet impactful, universities can foster continuous professional development without burdening participants' schedules or budgets.

Academic leaders, such as Heads of Department and Programme Coordinators, should integrate collegial behaviors into daily decision-making practices. The evidence from this training shows that these behaviors can be learned and are not solely driven by individual personality traits. Thus, leaders can create a more collaborative and supportive environment that encourages ongoing development in collegial practices.

6.3. Recommendations for Future Research

Looking forward, future studies should incorporate longitudinal follow-up to examine sustained behavioral change in the workplace, as understanding long-term impacts is critical for validating training effectiveness (Asaari et al., 2016). This would provide insights into how training influences academic professionals over time, guiding improvements in program design and delivery.

Utilizing observation or 360-degree feedback could enrich understanding of behavioral transfer beyond self-reported measures (Chiaburu et al., 2010). Such methodologies can provide a more comprehensive picture of how newly acquired skills are integrated into daily practices, establishing a multi-faceted evaluation approach that captures the intricacies of behavior change.

Additionally, conducting comparative analyses between disciplines (e.g., STEM vs. Social Sciences) may reveal variations in the practice of collegial leadership. Different academic fields may emphasize distinct aspects of collaboration and leadership; understanding these differences can inform tailored training programs that resonate with specific departmental cultures.

Finally, future research could examine organizational outcomes such as team climate, productivity, and conflict resolution following the training, providing a more comprehensive view of the training's broader impact on organizational performance (Ladyshevsky & Flavell, 2011). Seeking to understand the effects of training on overall institutional environments can guide strategic decisions that enhance faculty engagement and collaboration.

Acknowledgement

This paper is funded by Putra-IPM Grant (9731000), Faculty of Educational Studies, Universiti Putra Malaysia.

Data Availability Statement

The data supporting the findings of this study are not publicly available due to ethical considerations related to participant confidentiality and the small, context specific sample drawn from a single institution. Public deposition of the dataset could increase the risk of participant re identification. Anonymized data and analysis syntax may be made available by the corresponding author upon reasonable request, subject to institutional approval.

References

- Abbas, M., Sarwar, M., & Rehman, R. (2024). Impact of in-service training on people-focused and democratic leadership skills among educational managers of punjab. *Journal of Social Research Development*, 5(2), 96-107. <https://doi.org/10.53664/jsrd/05-02-2024-08-96-107>
- Alghanaim, N., Hart, J., & Finn, G. (2025). Scoping review: the effectiveness of interprofessional virtual reality simulation. *Journal of Medical Education and Curricular Development*, 12, 1-18. <https://doi.org/10.1177/23821205251346327>
- Almessabi, A. (2021). Culturally foreign teachers' perceptions of school climate and its relationship to their self-efficacy. *Sage Open*, 11(3). <https://doi.org/10.1177/21582440211043927>
- Amtu, O., Souisa, S., Joseph, L., & Lumamuly, P. (2021). Contribution of leadership, organizational commitment and organizational culture to improve the quality of higher education. *International Journal of Innovation*, 9(1), 131-157. <https://doi.org/10.5585/iji.v9i1.18582>
- Anumolu, R. and Jain, A. (2025). A comparative study of democratic (participative) and laissez-faire (delegative) leadership styles in enhancing employee retention in higher educational institutions across major cities of gujarat. *Int J Sci Res Humanities and Social Sciences*, 2(3), 145-155. <https://doi.org/10.32628/ijsrhss252315>
- Arabi, E. and Garza, T. (2022). Training design enhancement through training evaluation: effects on training transfer. *International Journal of Training and Development*, 27(2), 191-219. <https://doi.org/10.1111/ijtd.12295>
- Asaari, M., Dwivedi, A., Lawton, A., & Desa, N. (2016). Academic leadership and organizational commitment in public universities of malaysia. *European Scientific Journal Esj*, 12(16), 329. <https://doi.org/10.19044/esj.2016.v12n16p329>
- Ayeleke, R., North, N., Dunham, A., & Wallis, K. (2019). Impact of training and professional development on health management and leadership competence. *Journal of Health Organization and Management*, 33(4), 354-379. <https://doi.org/10.1108/jhom-11-2018-0338>
- Baldwin, T. and Ford, J. (1988). Transfer of training: a review and directions for future research. *Personnel Psychology*, 41(1), 63-105. <https://doi.org/10.1111/j.1744-6570.1988.tb00632.x>
- Beiting, K., Chippendale, R., Goroncy, A., & Duggan, M. (2023). Geriatrics fellows learning online and together (geri-a-float): a sustainable model of learning and support. *Journal of the American Geriatrics Society*, 71(9), 2902-2912. <https://doi.org/10.1111/jgs.18458>
- Boet, S., Bould, M., Fung, L., Qosa, H., Perrier, L., Tavares, W., ... & Tricco, A. (2014). Transfer of learning and patient outcome in simulated crisis resource management: a systematic review. *Canadian Journal of Anesthesia/Journal Canadien D Anesthésie*, 61(6), 571-582. <https://doi.org/10.1007/s12630-014-0143-8>

Cadena-Povea, H., Hernández-Martínez, M., Bastidas-Amador, G., & Torres-Andrade, H. (2025). What pushes university professors to burnout? a systematic review of sociodemographic and psychosocial determinants. *International Journal of Environmental Research and Public Health*, 22(8), 1214. <https://doi.org/10.3390/ijerph22081214>

Chauhan, R., Ghosh, P., & Kumar, S. (2024). Enhancing transfer of training in a power major in India. *The Journal of Management Development*, 44(1), 1-24. <https://doi.org/10.1108/jmd-07-2023-0212>

Chiaburu, D., Dam, K., & Hutchins, H. (2010). Social support in the workplace and training transfer: a longitudinal analysis. *International Journal of Selection and Assessment*, 18(2), 187-200. <https://doi.org/10.1111/j.1468-2389.2010.00500.x>

Cook, J., Money, B., Dyer, D., Whiteley, G., & Huntington, M. (2024). Living, leading & medicine: a two-tiered leadership development programme for family medicine residents. *The Clinical Teacher*, 21(5). <https://doi.org/10.1111/tct.13765>

Cowman, M. and McCarthy, A. (2016). The impact of demographic and situational factors on training transfer in a health care setting. *The Irish Journal of Management*, 35(2), 129-142. <https://doi.org/10.1515/ijm-2016-0009>

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Cromwell, S. and Kolb, J. (2004). An examination of work-environment support factors affecting transfer of supervisory skills training to the workplace. *Human Resource Development Quarterly*, 15(4), 449-471. <https://doi.org/10.1002/hrdq.1115>

Douphrate, D., Rodriguez, A., Kines, P., & Javid, A. (2024). Safety leadership training effectiveness evaluation on behavior change among large-herd u.s. dairy farm supervisors. *Workplace Health & Safety*, 73(4), 180-192. <https://doi.org/10.1177/21650799241302817>

Erdemir, B. (2023). Destructive leadership and organizational culture as the predecessors of academic mobbing. *Eğitim Bilim Ve Araştırma Dergisi*, 4(1), 171-220. <https://doi.org/10.54637/ebad.1259385>

Esterhazy, R., Lange, T., Bastiansen, S., & Wittek, A. (2021). Moving beyond peer review of teaching: a conceptual framework for collegial faculty development. *Review of Educational Research*, 91(2), 237-271. <https://doi.org/10.3102/0034654321990721>

Facteau, J. D., Dobbins, G. H., Russell, J. E., Ladd, R. T., & Kudisch, J. D. (1995). The influence of general perceptions of the training environment on pretraining motivation and perceived transfer of training. *Journal of Management*, 21(1), 1-25.

Farhadi, R., Azandehi, B., Amuei, F., Ahmadi, M., Zazoly, A., & Ghorbani, A. (2023). Enhancing residents' neonatal resuscitation competency through team-based simulation training: an intervention educational study. *BMC Medical Education*, 23(1). <https://doi.org/10.1186/s12909-023-04704-4>

Farzana, J., Khurum, S., & Alwi, K. (2022). Educators perception about leadership competencies: a comparison of public and private universities of pakistan. *Pakistan Languages and Humanities Review*, 6(1). [https://doi.org/10.47205/plhr.2022\(6-i\)6](https://doi.org/10.47205/plhr.2022(6-i)6)

Fernandez, C., Noble, C., Jensen, E., & Chapin, J. (2016). Improving leadership skills in physicians: a 6-month retrospective study. *Journal of Leadership Studies*, 9(4), 6-19. <https://doi.org/10.1002/jls.21420>

Grossman, R. and Salas, E. (2011). The transfer of training: what really matters. *International Journal of Training and Development*, 15(2), 103-120. <https://doi.org/10.1111/j.1468-2419.2011.00373.x>

Guo, Y. (2024). University climate affecting the teachers' motivation towards improvement of teachers' classroom performance in shandong qilu university of technology. *Pacific International Journal*, 6(4), 92-97. <https://doi.org/10.55014/pij.v6i4.482>

Hamid, M., Afzal, M., Khan, S., & Yasmeen, R. (2021). Evaluation of training course on “integrated management of neonatal & childhood illness” (imnci) using kirkpatrick model at level-2. *Health Professions Educator Journal*, 4(1), 14-18. <https://doi.org/10.53708/hpej.v4i1.100>

Hayes, M. and Ingram, K. (2019). Leadership and career development curriculum in australian dental and oral health programmes. *European Journal of Dental Education*, 23(3), 378-384. <https://doi.org/10.1111/eje.12444>

Holton III, E., Bates, R. A., Ruona, W. E., & Leimbach, M. (1998). Development and validation of a generalized learning transfer climate questionnaire: Final report. In R. Torraco (Eds.), *AHRD 1998 Conference Proceedings* (pp.482-489). Academic of Human Resource Development.

Hughes, A., Gregory, M., Joseph, D., Sonesh, S., Marlow, S., Lacerenza, C., ... & Salas, E. (2016). Saving lives: a meta-analysis of team training in healthcare. *Journal of Applied Psychology*, 101(9), 1266-1304. <https://doi.org/10.1037/apl0000120>

Jamali, A., Bhutto, A., Khaskhely, M., & Sethar, W. (2022). Impact of leadership styles on faculty performance: moderating role of organizational culture in higher education. *Management Science Letters*, 12(1), 1-20. <https://doi.org/10.5267/j.msl.2021.8.005>

Jones, W. and Dexter, S. (2014). How teachers learn: the roles of formal, informal, and independent learning. *Educational Technology Research and Development*, 62(3), 367-384. <https://doi.org/10.1007/s11423-014-9337-6>

Kelley, L., Broadfoot, K., McCreight, M., Wills, A., Leonard, C., Connelly, B., ... & Burke, R. (2023). Implementation and evaluation of a training curriculum for experienced nurses in care coordination. *Journal of Nursing Care Quality*, 38(3), 286-292. <https://doi.org/10.1097/ncq.0000000000000698>

Kontoghiorghes, C. (2008). Predicting motivation to learn and motivation to transfer learning back to the job in a service organization: a new systemic model for training effectiveness. *Performance Improvement Quarterly*, 15(3), 114-129. <https://doi.org/10.1111/j.1937-8327.2002.tb00259.x>

Lacerenza, C., Reyes, D., Marlow, S., Joseph, D., & Salas, E. (2017). Leadership training design, delivery, and implementation: a meta-analysis. *Journal of Applied Psychology*, 102(12), 1686-1718. <https://doi.org/10.1037/apl0000241>

Ladyshevsky, R. and Flavell, H. (2011). Transfer of training in an academic leadership development program for program coordinators. *Educational Management Administration & Leadership*, 40(1), 127-147. <https://doi.org/10.1177/1741143211420615>

Leránoz-Iglesias, M., Fernández-Morante, C., Cebreiro-López, B., & Pereira, C. (2023). Study on the Collaboration between University and Educational Centers Mentors in the Development of the In-School Education Placements in Official University Degrees Qualifying for the Teaching Profession: The Case of the University of Santiago de Compostela. *Education Sciences*, 13(2), 104. <https://doi.org/10.3390/educsci13020104>

Martin, H. (2010). Workplace climate and peer support as determinants of training transfer. *Human Resource Development Quarterly*, 21(1), 87-104. <https://doi.org/10.1002/hrdq.20038>

Maulidya, S., Larassati, S., Fajariya, U., Salwa, S., Maulana, M., & Kusumaningrum, H. (2024). Evaluasi pelatihan manajemen keorganisasian dalam program trailer dengan menggunakan model kirkpatrick. *Student Scientific Creativity Journal*, 3(1), 198-212. <https://doi.org/10.55606/ssci-amik.v3i1.4934>

McGowan, E., Hale, J., Bezner, J., Harwood, K., Green-Wilson, J., & Stokes, E. (2020). Leadership development of health and social care professionals: a systematic review. *BMJ Leader*, 4(4), 231-238. <https://doi.org/10.1136/leader-2020-000211>

Merwe, W., Maree, C., Yazbek, M., & Cochrane, M. (2025). Empowered to lead: the effect of mentoring and tutoring on peer tutor leadership development. *New Directions for Student Leadership*, 2025(186), 75-80. <https://doi.org/10.1002/yd.20678>

Miller, D., Crandall, C., Washington, C., & McLaughlin, S. (2012). Improving teamwork and communication in trauma care through in situ simulations. *Academic Emergency Medicine*, 19(5), 608-612. <https://doi.org/10.1111/j.1553-2712.2012.01354.x>

Namutebi, A. (2024). Development and evaluation of leadership training and development programs in uganda. *International Journal of Leadership and Governance*, 4(1), 53-63. <https://doi.org/10.47604/ijlg.2413>

Nengyanti, N., Yusnaini, Y., Imania, K., & Santoso, A. (2024). Policy Transfer for Sexual Violence Prevention and Management in Indonesian Higher Education Institutions. *Multidisciplinary Journal of Gender Studies*, 13(13), 137-155. <https://doi.org/10.17583/generos.12738>

Ng, K. (2014). Supervisory practices and training transfer: lessons from malaysia. *Asia Pacific Journal of Human Resources*, 53(2), 221-240. <https://doi.org/10.1111/1744-7941.12044>

Ng, P., Chow, P., Wong, P., & Luk, W. (2024). The impact of digital technology training on developing academics' digital competence in higher education context. *Education + Training*, 66(9), 1276-1292. <https://doi.org/10.1108/et-10-2023-0417>

Phillips, D., Bhojedat, J., Phillips, S., Henry, C., & Stewart-Fox, T. (2024). Exploring the readiness of schools for distributed leadership: perspectives of private school teachers in Guyana. *Creative Education*, 15(02), 249-277. <https://doi.org/10.4236/ce.2024.152015>

Saks, A. and Burke, L. (2012). An investigation into the relationship between training evaluation and the transfer of training. *International Journal of Training and Development*, 16(2), 118-127. <https://doi.org/10.1111/j.1468-2419.2011.00397.x>

Saidi, Z. A. & Abd Aziz, M. F. (2025). Empowering student leadership in higher education: A literature review on key competencies for effective leadership. *Journal of Institutional Research South East Asia*, 23(1), 376-399.

Saidi, Z.A., Asri, A.S., & Abd Aziz, M.F. (2025). Exploring the relationship between participation, needs, readiness, and planning in university student volunteering programs. *Journal of Institutional Research South East Asia*, 23(1), 401-417.

Schaller, M., Gencheva, M., Gunther, M., & Weed, S. (2023). Training doctoral students in critical thinking and experimental design using problem-based learning. *BMC Medical Education*, 23(1). <https://doi.org/10.1186/s12909-023-04569-7>

Seigel, M. L., & Miner-Rubino, K. (2009). Measuring the value of collegiality among law professors. *Faulkner L. Rev.*, 1, 257-287.

Sharif, S., Braimah, M., & Dogbey, A. (2022). Academic supports, motivation to learn, motivation to transfer and transfer of training: a comparative analysis of public and private universities. *European Journal of Training and Development*, 47(5/6), 507-532. <https://doi.org/10.1108/ejtd-05-2021-0068>

Sim, J. and Radloff, A. (2008). Enhancing reflective practice through online learning: impact on clinical practice. *Biomedical Imaging and Intervention Journal*, 4(1). <https://doi.org/10.2349/biij.4.1.e8>

Singh, S. and Patel, A. (2025). A study on the relationship between employee engagement and faculty retention in higher education institutions. *International Journal of Management and Humanities*, 11(11), 1-5. <https://doi.org/10.35940/ijmh.h1815.11110725>

Sparr, J., Knipfer, K., & Willems, F. (2016). How leaders can get the most out of formal training: the significance of feedback-seeking and reflection as informal learning behaviors. *Human Resource Development Quarterly*, 28(1), 29-54. <https://doi.org/10.1002/hrdq.21263>

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Pearson/Allyn & Bacon.

Tan, W. (2024). Empowering examiners to develop doctoral assessment literacy: a situated learning perspective. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1345661>

Tapic, J. and Baguio, J. (2024). Restoration leadership practices and collegial model of public elementary school teachers. *Asian Journal of Education and Social Studies*, 50(12), 357-365. <https://doi.org/10.9734/ajess/2024/v50i121704>

Walter, S., Schmutz, J., & Grote, G. (2024). Train one and get a crew: a multilevel evaluation of team-centric leadership training for airport security screening crew leaders. *Team Performance Management*, 30(7/8), 195-215. <https://doi.org/10.1108/tpm-06-2023-0047>

Wennerberg, J. and McGrath, C. (2022). Breaking the isolation: a study of university teachers' collective development. *Journal of Praxis in Higher Education*, 4(1), 7-27. <https://doi.org/10.47989/kpdc110>

Wißhak, S., Schäfer, P., & Wavere, L. (2025). Train-The-Trainer: A Generic Offer-And-Use Model for the Development of Trainers. *International Journal of Training and Development*, 29(4), 449-461. <https://doi.org/10.1111/ijtd.12370>

Winebrake, K., McDonough, R., & Mara, M. (2024). Teaching students leadership knowledge and skills: a 10-year comparison of predoctoral dental school curricula. *Journal of Dental Education*, 89(1), 98-106. <https://doi.org/10.1002/jdd.13687>

Yaghi, A. and Bates, R. (2020). The role of supervisor and peer support in training transfer in institutions of higher education. *International Journal of Training and Development*, 24(2), 89-104. <https://doi.org/10.1111/ijtd.12173>

Young, J., Fawcett, K., & Gillman, L. (2022). Evaluation of an immersive simulation programme for mental health clinicians to address aggression, violence, and clinical deterioration. *International Journal of Mental Health Nursing*, 31(6), 1417-1426. <https://doi.org/10.1111/inm.13040>

Zulkifly, N. (2023). Academic collegial leadership training transfer: an integrated evaluation approach. *International Journal of Academic Research in Progressive Education and Development*, 12(4), 313-329. <https://doi.org/10.6007/ijarped/v12-i4/19675>

Zulkifly, N., Ismail, I., Asimiran, S., & Alias, S. (2021). Factors influencing transfer of training among academics in public universities in malaysia. *International Journal of Academic Research in Progressive Education and Development*, 10(2), 184-200. <https://doi.org/10.6007/ijarped/v10-i2/9686>

Appendix A

Collegial Leadership Training Programme Curriculum

The Collegial Leadership Training Programme (CLTP) was designed as a structured professional development intervention aimed at enhancing collegial leadership competencies among university academics. The programme was delivered in a single training cycle and comprised four integrated sessions conducted within one training day, with a total contact time of eight hours.

Session 1: Introduction to Collegial Leadership
Duration: 1.5 hours

This session introduced the concept of collegial leadership in higher education, focusing on shared decision making, professional respect, and collaborative academic practice. Participants engaged in guided discussion on common collegial challenges in academic departments and reflected on their own leadership experiences.

Session 2 Communication and Constructive Dialogue
Duration: 2 hours

This session focused on interpersonal communication skills, including active listening, constructive feedback, and respectful disagreement in academic contexts. Case-based scenarios drawn from typical university settings were used to facilitate discussion and peer learning.

Session 3: Collaborative Problem Solving and Conflict Management
Duration: 2.5 hours

Participants worked in small groups to analyse realistic academic leadership cases involving conflict, role ambiguity, and shared responsibility. The session emphasised collaborative decision making, perspective taking, and collegial resolution strategies.

Session 4: Reflective Practice and Application to Academic Roles
Duration: 2 hours

The final session focused on reflective dialogue and application of collegial leadership principles to participants' own academic roles. Participants developed individual action reflections outlining how collegial behaviours could be enacted within their departments following the training.

Across all sessions, active learning strategies were prioritised, including group discussion, case analysis, and guided reflection, to encourage behavioural practice aligned with collegial leadership principles.

Appendix B

Training Delivery and Fidelity Assurance

To ensure consistency and fidelity in programme delivery, the CLTP followed a standardised training outline specifying session objectives, content coverage, time allocation, and learning activities. All sessions were delivered according to the planned schedule, and no deviations from the core curriculum were reported. The programme was facilitated by members of the research team with expertise in academic leadership, human resource development, and training evaluation. Facilitators had prior experience delivering leadership and professional development programmes in higher education settings. Their role was to guide discussion, facilitate case-based activities, and ensure alignment between session activities and intended learning outcomes.

Fidelity was supported procedurally through the use of common training materials, identical case scenarios across groups, and consistent facilitation approaches across sessions. Although formal observer-based fidelity checklists were not employed, adherence to the planned curriculum was maintained through structured session guides and facilitator coordination before and during programme delivery. This approach ensured that all participants were exposed to comparable learning experiences, supporting the internal consistency of the training intervention evaluated in this study.

Attendance records were maintained throughout the Collegial Leadership Training Programme to ensure full exposure to the intervention. All 30 participants attended the complete training programme, with no attrition observed across sessions. As the programme was delivered within a single training cycle, full attendance ensured that all participants received identical content, learning activities, and contact hours. The facilitators comprised senior academic staff with backgrounds in educational leadership, professional development, and training transfer research. All facilitators held doctoral qualifications and had prior experience in designing and delivering leadership development programmes for university academics.

Appendix C

Sample Case-Study

Case Title

Navigating Collegial Decision Making in a Teaching Allocation Dispute

Case Context

Dr A is a senior academic in a public university department and has recently been appointed as a programme coordinator. The department comprises academics with varying levels of seniority, disciplinary backgrounds, and administrative responsibilities. While the department values collegiality and shared decision making, tensions occasionally arise due to workload distribution and competing priorities.

Scenario

At the beginning of the semester, the department faces a shortage of lecturers for a core undergraduate course due to an unexpected staff transfer. As programme coordinator, Dr A proposes reallocating teaching responsibilities to ensure the course can proceed smoothly. The proposed allocation requires two senior academics to take on an additional tutorial group. During the departmental meeting, one senior academic expresses dissatisfaction, arguing that teaching allocations should prioritise research active staff and that the decision was made without sufficient consultation. Another colleague supports Dr A's proposal but remains hesitant to speak openly. The discussion becomes tense, and several members disengage from the conversation. Dr A must decide how to respond in a way that maintains collegial relationships while ensuring that departmental responsibilities are met.

Discussion

Participants were asked to discuss the following questions in small groups:

1. How should Dr A facilitate the discussion to ensure all voices are heard while maintaining a respectful and constructive tone?
2. What collegial leadership behaviours are relevant in managing disagreement in this situation?
3. How can shared decision making be practised without delaying necessary action?
4. What strategies could help rebuild trust and collaboration following the disagreement?
5. How might this situation be handled differently to prevent similar issues in the future?

Learning Focus

This case was designed to encourage participants to practise collegial leadership behaviours, including constructive dialogue, perspective taking, collaborative problem solving, and conflict management within a realistic academic context. The scenario reflects common challenges faced by academic leaders and allows participants to rehearse responses that balance task requirements with relational considerations.

Appendix D

Instruments and Item-level Statistics

Collegiality Leadership

	1	2	3	4	5
1. I honor agreements.	1	2	3	4	5
2. I “pull my own weight.”	1	2	3	4	5
3. I am willing to help each other.	1	2	3	4	5
4. I comment on each other’s work.	1	2	3	4	5
5. I am encouraging and empowering.	1	2	3	4	5
6. I create a cooperative and supportive environment.	1	2	3	4	5
7. I initiate and participate in informal conversations about the organizations (about teaching, research, etc.)	1	2	3	4	5
8. I share product of my own effort with each other.	1	2	3	4	5
9. I work to foster harmony.	1	2	3	4	5
10. I take on special assignments (given to me by head of department, deputy dean or dean).	1	2	3	4	5
11. I participate in all aspects of organizational life (e.g., attend functions).	1	2	3	4	5

1-Strongly disagree; 2-Disagree; 3-Somewhat agree; 4-Agree; 5-Strongly Agree

Item-Level Descriptive Statistics for Collegial Leadership Scale at Pre- and Post-Training

Item	Pre Mean	Pre SD	Post Mean	Post SD
I honor agreements.	4.50	0.63	4.77	0.43
I pull my own weight.	4.47	0.63	4.73	0.45
I am willing to help each other.	4.70	0.47	4.80	0.41
I comment on each other’s work.	3.60	0.77	4.23	0.68
I am encouraging and empowering.	4.43	0.57	4.60	0.50
I create a cooperative and supportive environment.	4.37	0.67	4.57	0.57
I initiate and participate in informal conversations about the organization (about teaching, research, etc.).	4.07	0.69	4.53	0.63
I share product of my own effort with each other.	4.17	0.79	4.30	0.70
I work to foster harmony.	4.37	0.61	4.60	0.50
I take on special assignments (given to me by head of	4.20	0.85	4.50	0.57

department, deputy dean or dean).

I participate in all aspects of organizational life (e.g., attend functions).

3.83

0.83

4.33

0.71

Training Transfer

	1	2	3	4	5
1. Peers have told me that my behavior has improved following a training course.	1	2	3	4	5
2. The productivity of my peers has improved due to the skills that I learned from training courses.	1	2	3	4	5
3. Absenteeism in my group has decreased due to the skills that I developed from training courses.	1	2	3	4	5
4. Turnover intention in my group has decreased due to the skills that I developed from training courses.	1	2	3	4	5
5. Morale of my work group is higher due to the skills that I developed from training courses.	1	2	3	4	5
6. My peers are more committed to the mission of the university due to the skills that I developed from training courses.	1	2	3	4	5
7. I am able to transfer the skills learned from training courses back to my actual job.	1	2	3	4	5
8. I have changed my job behavior in order to be consistent with the material taught in training courses.	1	2	3	4	5
9. My actual job performance has improved due to the skills that I learned from training courses.	1	2	3	4	5

1-Strongly disagree; 2-Disagree; 3-Somewhat agree; 4-Agree; 5-Strongly Agree

Item-Level Descriptive Statistics for Training Transfer Scale (Post-Training)

Item	Mean	SD
Peers have told me that my behavior has improved following a training course.	4.10	0.71
The productivity of my peers has improved due to the skills that I learned in training courses.	4.17	0.59
Absenteeism in my group has decreased due to the skills that I developed in training courses.	3.90	0.61
Turnover intention in my group has decreased due to the skills that I developed in training courses.	3.87	0.68

Morale of my work group is higher due to the skills that I developed in training courses.	4.23	0.68
My peers are more committed to the mission of the government due to the skills that I developed in training courses.	4.20	0.71
I am able to transfer the skills learned in training courses back to my actual job.	4.57	0.57
I have changed my job behavior in order to be consistent with the material taught in training courses.	4.33	0.55
My actual job performance has improved due to the skills that I learned in training courses.	4.40	0.67

Competency-Based Framework for Classroom Assessment Practices in TVET: The Role of Technical, Industry 4.0, and Career Adaptability Skills

ABSTRACT

Authors: Siti Raudhah M. Yusop¹, Nurhidayah Yaakop^{2*} and Nor Azlan Ahmad³

Affiliation: ¹Technical Vocational Education and Training Department, Ministry of Education, Putrajaya, Malaysia

²Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

³Faculty of Science and Information Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

*Corresponding author:
nurhidayahyaakop@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Technical and Vocational Education and Training (TVET) play a critical role in equipping the workforce with specialized skills aligned with evolving industry demands. Central to this is effective Classroom Assessment (CA), which is essential for monitoring progress, providing feedback, and fostering career adaptability among students. However, TVET CA practices often face challenges such as limited teacher understanding, insufficient resources, and time constraints. This study investigates teacher perceptions regarding how classroom assessment supports the development of students' Technical Skills, Industry 4.0 Generic Skills, and Career Adaptability Skills. A competency-based framework was developed and tested using Partial Least Squares Structural Equation Modelling (PLS-SEM) with responses from 418 senior TVET teachers (Heads of Programme) across Malaysian Vocational Colleges. The analysis evaluated reliability, convergent and discriminant validity, predictive accuracy, and mediation effects. Findings reveal that Technical Skills ($\beta = 0.268$, $t = 5.039$) and Career Adaptability Skills ($\beta = 0.389$, $t = 5.007$) have significant positive effects on CA Practices. While Industry 4.0 Generic Skills showed a marginal direct effect on CA practices ($\beta = 0.119$, $t = 1.467$, $p < 0.10$), they exert a significant indirect influence through the mediation of Career Adaptability Skills ($\beta = 0.328$, $t = 4.934$, $p < 0.01$). These insights suggest that optimizing classroom assessment can significantly strengthen skill development and graduate employability. The findings offer practical implications for trainers, policymakers, and curriculum developers in enhancing teaching practices through effective assessment strategies.

Keywords: Technical Skills, Industry 4.0 Generic Skills, Career Adaptability Skills, Classroom Assessment Practices, TVET, PLS-SEM.

1. Introduction

UNESCO defines Technical and Vocational Education and Training (TVET) as education, training and skills development across various occupational fields, including production, services and livelihoods (UNESCO-UNEVOC, 2017). As a key component of lifelong learning, TVET can be delivered at the secondary, post-secondary and tertiary levels, encompassing work-based learning, continuing education and professional development pathways that lead to recognised qualifications. In addition, TVET plays a critical role in achieving the United Nations Sustainable Development Goals (SDGs), particularly in increasing the number of youth and adults equipped with relevant technical and vocational skills to support employment, decent work and entrepreneurship (United Nation, 2020). The SDGs agenda further emphasises the importance of promoting full and productive employment, as well as sustained, inclusive and sustainable economic growth (UNESCO, 2021). In this context, expanding the scope of TVET, strengthening delivery mechanisms, preparing competent trainers and enabling the workforce to adapt to technological advancements are essential strategies for supporting the transition towards a green and knowledge-based economy.

In Malaysia, educational transformation initiatives such as the Malaysia Education Development Plan 2013–2025 highlight the need for a more integrated and holistic approach to student assessment. The Development of an Integrated Assessment System Methodology introduced by the Ministry of Education underscores the importance of evaluating students' knowledge, skills and competencies in a balanced and comprehensive manner (Kementerian Pendidikan Malaysia, 2018). However, assessment practices in Malaysia continue to be largely influenced by centralised and examination-oriented approaches (Adam et al. 2019; Raman & Yamat, 2014; Ratnam-Lim & Tan, 2015). This reliance on examination-based systems has been widely debated among educators, policymakers and the academic community, as such approaches often prioritise cognitive outcomes while overlooking practical skills, values and work-related competencies (Azmanirah Ab Rahman et al. 2020; Rahman et al. 2019).

Within the context of TVET, classroom assessment (CA) plays a pivotal role in evaluating students' competencies and skills through both theoretical and practical learning activities. Nevertheless, inadequately designed assessment practices may negatively affect the employability of TVET graduates. One of the key challenges is that existing CA practices may not sufficiently equip students with real-world skills and knowledge required in the labour market (Winther & Klotz, 2013). For instance, assessments that focus predominantly on theoretical knowledge rather than practical performance may result in graduates who are not adequately prepared to meet technical job requirements (Ahmad Zaidi Husin, 2019). This highlights the need to re-examine current assessment practices to ensure closer alignment with industry expectations.

Accordingly, the development of assessment methodologies, frameworks and instruments must be aligned with industry demands to support the implementation of industry-driven curricula. Assessment practices that fail to reflect industry expectations may result in graduates who lack the required competencies for employment (Rina Febriana et al. 2017). Therefore, curriculum enhancement efforts must be supported by competency-based assessment approaches that effectively measure knowledge, skills and competencies relevant to current industry needs. Such improvements are essential not only for enhancing graduate employability but also for reducing reliance on foreign skilled labour and addressing skill mismatches among TVET graduates (Campbell, 2016).

Despite these developments, limited empirical studies have examined how multiple competency domains—particularly technical skills, Industry 4.0 generic skills and career adaptability skills—interact within the context of classroom assessment practices in TVET. Existing studies have largely focused on individual competencies or employability outcomes, with less emphasis on understanding the structural relationships among these constructs (Mtshali & Ramaligela, 2020; Rasul et al. 2010). Furthermore, many previous studies rely primarily on student self-reports, with limited attention given to teachers' perspectives as primary evaluators of students' competencies in classroom settings (Glogger-Frey et al. 2018; Maisiri et al. 2021; Tan et al. 2017). Given that teachers play a central role

in observing, assessing and reporting students' performance, their professional judgments provide valuable insights into students' competency development.

Therefore, this study aims to examine the relationships between technical skills, Industry 4.0 generic skills, career adaptability skills and classroom assessment practices from the perspective of TVET teachers. Specifically, this study investigates how these competencies influence classroom assessment practices, as well as the mediating role of career adaptability skills in these relationships. By adopting a structural modelling approach, this study contributes to the development of a more comprehensive understanding of competency-based assessment in TVET and provides empirical evidence to support the alignment between curriculum, assessment practices and industry requirements.

Specifically, this study aims:

RQ1: To what extent do teachers perceive students' Technical Skills as influencing their Career Adaptability Skills?

RQ2: To what extent do teachers perceive students' Industry 4.0 Generic Skills as influencing their Career Adaptability Skills?

RQ3: How do teachers perceive the influence of students' Career Adaptability Skills on TVET CA Practices?

RQ4: How do teachers perceive the influence of students' Technical Skills on TVET CA Practices?

RQ5: How do teachers perceive the influence of students' Industry 4.0 Generic Skills on TVET CA Practices?

RQ6: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Technical Skills and CA Practices?

RQ7: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Industry 4.0 Generic Skills and CA Practices?

The proposed initial model, as presented in Figure 1, is developed based on an extensive literature review. The following hypotheses (Ha) are to evaluate the overall model.

Ha1: Teachers perceive that students' Technical Skills have a significant positive influence on their Career Adaptability Skills.

Ha2: Teachers perceive that students' Industry 4.0 Generic Skills have a significant positive influence on their Career Adaptability Skills.

Ha3: Teachers perceive that students' Career Adaptability Skills have a significant positive influence on TVET CA Practices.

Ha4: Teachers perceive that students' Technical Skills have a significant positive influence on TVET CA Practices.

Ha5: Teachers perceive that students' Industry 4.0 Generic Skills have a significant positive influence on TVET CA Practices.

Ha6: Teachers perceive that Career Adaptability Skills significantly mediate the relationship between Technical Skills and TVET CA Practices.

Ha7: Teachers perceive that Career Adaptability Skills significantly mediate the relationship between Industry 4.0 Generic Skills and TVET CA Practices.

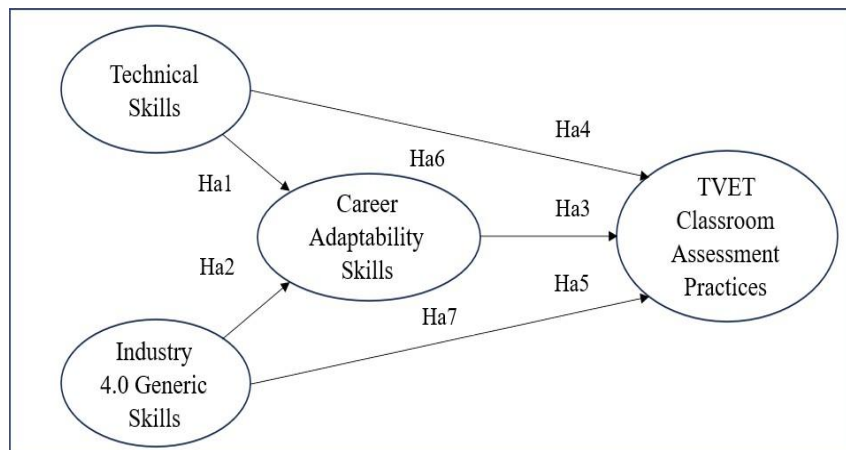


Figure 1. The proposed initial measurement model

Note: Representation of the two exogenous constructs (Technical Skills, Industry 4.0 Generic Skills), the mediator (Career Adaptability Skills), and the endogenous constructs (TVET Classroom Assessment Practices).

2. Literature Review

Technical and Vocational Education and Training (TVET) has become a critical educational agenda in developing highly skilled human capital to meet the evolving demands of Industry 4.0 and support Malaysia's aspiration to become a developed nation. The increasing demand for skilled labour is evident, as the government anticipates the creation of up to 1.5 million new jobs, of which approximately 60% will require skilled workers. However, only about 31% of these positions have been filled, falling short of the targeted 35% and resulting in a 4% skills gap (Makhbul & Latif, 2019). This mismatch highlights a significant challenge in aligning workforce competencies with industry requirements.

Despite continuous efforts to strengthen TVET, concerns remain regarding the misalignment between graduates' competencies and rapid technological advancements. Previous studies indicate that Malaysian employees' skills are still not fully aligned with the needs of Industry 4.0 (Aizuddin Saari et al. 2021). Furthermore, insufficient collaboration between industry and educational institutions has limited the effectiveness of curriculum enhancement, contributing to lower competency levels among TVET students (Sohimi et al. 2019). Empirical evidence also supports this concern, as a study involving 136 students across seven TVET learning domains at the Technical University of Malaysia revealed that students' readiness for Industry 4.0 skills remains inadequate (Afferro Ismail et al. 2020). These findings suggest the need for a more structured and industry-aligned approach to competency development within TVET systems.

In response to these challenges, competency-based frameworks such as the COMET model have been introduced to systematically guide the development and assessment of vocational competencies. The COMET competency model, developed by Rauner et al. (2013), provides a comprehensive framework that encompasses six key elements: competence, performance domains, performance statements, levels of mastery, assessment methods and development activities. This model enables educators to define essential capabilities, evaluate observable performance and design appropriate assessment strategies that reflect real workplace requirements. The integration of the COMET model into TVET programmes has been shown to improve the alignment between graduate competencies and industry expectations, thereby enhancing students' readiness and employability in Industry 4.0 contexts (Lahn & Nore, 2019).

The importance of competency development is further reflected in graduate employability outcomes. Data from the Ministry of Higher Education indicate relatively high employability rates among Malaysian TVET graduates, with Malaysian Technical University Network (MTUN) achieving 87.1%, polytechnics 88.6% and community colleges 97.2%. Similarly, findings from the Mersing Polytechnic

Graduate Tracking Study show that 91% of graduates in 2016 and 86.3% in 2017 secured employment, although the percentage dropped to 41.5% in 2018, with an increasing number of graduates working outside their field of study (Nor Zamira Othman & Siti Farah Hussin, 2019). While these statistics demonstrate positive outcomes, they also indicate potential issues related to job mismatch, suggesting that employability should not only be measured in terms of employment rates but also alignment with relevant competencies. Therefore, strengthening competency-based education and assessment remains essential to ensure that graduates are adequately prepared for industry demands.

2.1. Technical Skills in TVET Education

Technical skills represent the core competencies required for effective performance in vocational and technical occupations. Within the context of TVET, these skills encompass practical knowledge, procedural abilities and specialised expertise that enable individuals to perform job-related tasks efficiently (Mohd Nur Hafiz Fauzi, 2017). Such competencies are typically developed through structured training, hands-on learning and workplace-based experiences, which are fundamental components of vocational education. The role of TVET institutions in developing technical competencies is particularly significant, as these institutions serve as the primary platform for preparing students to meet labour market requirements.

Technical skills are essential in sectors that require operational precision and specialised knowledge, especially in industries undergoing rapid technological transformation. As automation and advanced technologies become more prevalent, the demand for technically competent workers continues to increase. Moreover, technical skills are strongly associated with employability and workplace performance. Empirical studies suggest that graduates with strong technical competencies are more likely to secure employment and adapt effectively to job requirements (Rina Febriana et al., 2017). However, the effectiveness of technical skill development largely depends on how these competencies are assessed within the educational system. This highlights the importance of aligning technical skill development with appropriate assessment practices to ensure that students are adequately prepared for real-world applications.

2.2. Industry 4.0 Generic Skills

The emergence of Industry 4.0 has significantly transformed the nature of work and the competencies required in modern industries. Industry 4.0 refers to the integration of digital technologies, artificial intelligence, automation and cyber-physical systems into industrial processes (Asnidatul Adilah Ismail & Razali Hassan, 2019). Consequently, the workforce is now expected to possess not only technical expertise but also a range of generic skills that enable them to function effectively in complex and technology-driven environments. Industry 4.0 generic skills include critical thinking, problem-solving, digital literacy, communication, teamwork and adaptability (Teo et al., 2021). These competencies are essential for managing complex tasks, collaborating across disciplines and responding to rapid technological changes.

In this context, technical skills alone are no longer sufficient; instead, a combination of technical and generic competencies is required to ensure workforce readiness. Educational institutions, particularly TVET providers, have increasingly emphasised the integration of Industry 4.0 competencies into their curricula (Khan et al., 2021). This integration aims to ensure that graduates remain relevant and competitive in the evolving labour market. Furthermore, these generic skills are closely linked to individuals' ability to adapt to career transitions and emerging occupational roles, suggesting a strong relationship between Industry 4.0 competencies and broader career-related outcomes.

2.3. Career Adaptability Skills

Career adaptability refers to an individual's readiness and capability to cope with current and anticipated career-related tasks, transitions and challenges. This concept is grounded in career construction theory, which emphasises the importance of adaptability in navigating dynamic and uncertain labour market conditions. Career adaptability is typically conceptualised through four dimensions: concern, control, curiosity and confidence (Savickas & Porfeli, 2012). These dimensions reflect an individual's ability to plan, make informed decisions, explore career opportunities and overcome challenges.

In the context of TVET, career adaptability is particularly important, as graduates often enter industries characterised by rapid technological changes and evolving job roles. Existing literature suggests that both technical skills and generic competencies play a significant role in enhancing career adaptability (Venugopal & Kumar, 2022). Individuals who possess strong technical knowledge and transferable skills are more likely to demonstrate higher levels of confidence, flexibility and resilience in managing their career pathways. Therefore, understanding how different competency domains contribute to career adaptability is essential for preparing students for long-term career success.

2.4. Classroom Assessment Practices in TVET

Classroom assessment practices are fundamental in evaluating students' learning outcomes and competency development in TVET education. Unlike traditional academic settings, assessment in TVET emphasises competency-based approaches that focus on students' ability to apply knowledge and skills in practical and authentic contexts. Effective classroom assessment involves a combination of formative and summative strategies (Brookhart, 2003; Kang & Furtak, 2021). Formative assessment provides continuous feedback to support learning, while summative assessment evaluates overall competency achievement. In TVET settings, assessment methods often include performance-based tasks, practical demonstrations, project-based learning and competency evaluations aligned with industry standards (Lewkowicz & Leung, 2021). Importantly, classroom assessment practices play a crucial role in shaping how students develop and demonstrate various competencies, including technical skills, Industry 4.0 generic skills and career adaptability. As teachers are directly involved in observing and evaluating students' performance during learning activities, their assessments provide valuable insights into students' competency development (Glogger-Frey et al., 2018). Therefore, examining the relationships between these competencies and classroom assessment practices is essential for improving teaching effectiveness and ensuring that TVET graduates meet industry expectations.

Based on the preceding discussion, it is evident that technical skills, Industry 4.0 generic skills and career adaptability skills represent critical competency domains required in contemporary TVET education. These competencies are not only essential for enhancing students' employability but are also closely associated with how learning outcomes are assessed within classroom settings (Ewing, 2017). Classroom assessment practices play a vital role in shaping and evaluating students' competency development through teachers' observations and professional judgments. Therefore, this study examines the relationships between technical skills, Industry 4.0 generic skills, career adaptability skills and classroom assessment practices from the perspective of TVET teachers, leading to the development of the following hypotheses.

2.5. Hypothesis Development

Based on the reviewed literature, technical skills, Industry 4.0 generic skills and career adaptability skills represent key competency domains required in Technical and Vocational Education and Training (TVET) (Hairi et al., 2019; Surono et al., 2020). These competencies are not only essential for enhancing students' employability but are also closely related to how learning outcomes are evaluated through classroom assessment practices. In TVET settings, teachers play a central role in observing, evaluating and reporting students' competencies during teaching and learning activities (Griffin et al., 2007). Therefore, this study examines the relationships among these constructs based on teachers' perceptions of students' competencies.

2.5.1. Technical Skills and Career Adaptability Skills

Technical skills constitute the foundational competencies required for effective performance in vocational and technical occupations. These skills enable individuals to perform tasks, solve technical problems and respond to workplace demands. In TVET learning environments, teachers continuously observe students' ability to demonstrate technical competencies during practical activities and assessments. From the perspective of competency development, technical skills may contribute to the development of career adaptability (Miyazaki et al., 2021). Students who demonstrate stronger technical competencies are more likely to exhibit confidence, control and flexibility in handling tasks and challenges. These attributes are closely aligned with the dimensions of career adaptability, including concern, control, curiosity and confidence (Kirchknopf, 2020). Therefore, based on teachers' observations, students with higher levels of technical skills are expected to demonstrate stronger career adaptability.

Ha1: Teachers perceive that students' Technical Skills have a significant positive influence on their Career Adaptability Skills.

2.5.2. Industry 4.0 Generic Skills and Career Adaptability Skills

Industry 4.0 generic skills refer to a set of transferable competencies, including critical thinking, problem-solving, adaptability, communication and collaboration, which are essential in technology-driven environments (Khan et al., 2021). These skills enable individuals to cope with uncertainty, manage complex tasks and adapt to changing work conditions. Within TVET contexts, teachers assess students' ability to demonstrate these competencies through various learning and assessment activities. Students who exhibit stronger Industry 4.0 generic skills are more likely to demonstrate adaptability in responding to new situations, exploring opportunities and managing career-related challenges (Chukwu et al., 2020). As such, these competencies are expected to contribute to the development of career adaptability.

Ha2: Teachers perceive that students' Industry 4.0 Generic Skills have a significant positive influence on their Career Adaptability Skills.

2.5.3. Career Adaptability Skills and TVET Classroom Assessment Practices

Career adaptability reflects students' readiness and capability to engage with learning tasks, respond to feedback and perform effectively in different assessment contexts. In TVET education, classroom assessment practices involve evaluating students' competencies through practical tasks, demonstrations and performance-based activities. Teachers play a key role in assessing how students apply their skills, manage tasks and respond to assessment requirements. Students who demonstrate higher levels of career adaptability are more likely to show confidence, persistence and flexibility in completing assessment tasks (Savickas & Porfeli, 2012). Therefore, from teachers' perspectives, higher levels of career adaptability are expected to be associated with better performance in classroom assessment practices.

Ha3: Teachers perceive that students' Career Adaptability Skills have a significant positive influence on TVET CA Practices.

2.5.4. Technical Skills and TVET Classroom Assessment Practices

Technical skills are directly evaluated in TVET classroom assessment practices through competency-based assessments, practical tasks and performance demonstrations. Teachers assess students' ability to apply technical knowledge and perform specific job-related tasks in structured learning environments (Glogger-Frey et al., 2018). Students who demonstrate stronger technical skills are more likely to

perform effectively in assessment activities, as these skills are closely aligned with the requirements of competency-based evaluation (Tun et al., 2021). Therefore, based on teachers' observations, higher levels of technical skills are expected to be positively associated with classroom assessment practices.

Ha4: Teachers perceive that students' Technical Skills have a significant positive influence on TVET CA Practices.

2.5.5. Industry 4.0 Generic Skills and TVET Classroom Assessment Practices

In addition to technical competencies, classroom assessment practices increasingly incorporate the evaluation of Industry 4.0 generic skills, such as problem-solving, teamwork and communication. These competencies are often assessed through project-based learning, collaborative activities and performance tasks. Teachers evaluate students' ability to demonstrate these skills during assessment activities (Hagos Baraki et al., 2016). Students who exhibit stronger Industry 4.0 generic skills are more likely to engage effectively in assessment tasks and demonstrate better overall performance (Ana et al., 2020). Therefore, it is expected that Industry 4.0 generic skills are positively related to classroom assessment practices.

Ha5: Teachers perceive that students' Industry 4.0 Generic Skills have a significant positive influence on TVET CA Practices.

2.5.6. The Mediating Role of Career Adaptability Skills

Career adaptability may also function as a mediating mechanism in the relationships between competency domains and classroom assessment practices. From a theoretical perspective, both technical skills and Industry 4.0 generic skills contribute to students' ability to adapt to learning and assessment environments (Savickas & Porfeli, 2012). Students with stronger technical and generic competencies are more likely to develop confidence, flexibility and control in performing tasks, which enhances their adaptability (de Guzman & Choi, 2013). This increased adaptability may subsequently influence how students perform in classroom assessment practices, as perceived by teachers. Accordingly, career adaptability is expected to mediate the relationships between technical skills, Industry 4.0 generic skills and classroom assessment practices.

Ha6: Teachers perceive that Career Adaptability Skills significantly mediate the relationship between Technical Skills and TVET CA Practices.

Ha7: Teachers perceive that Career Adaptability Skills significantly mediate the relationship between Industry 4.0 Generic Skills and TVET CA Practices.

Based on the reviewed literature, it is evident that technical skills, Industry 4.0 generic skills and career adaptability skills are interrelated competencies that influence students' performance in TVET learning environments. As teachers play a central role in assessing students' competencies through classroom assessment practices, their perceptions provide valuable insights into students' competency development. Therefore, this study proposes a structural model to examine the relationships among these constructs, leading to the following hypotheses.

3. Methodology

This study conducted two types of data analysis, namely descriptive analysis and inferential analysis. The survey instruments were administered using both Google Forms and printed copies. Data collection was conducted over a period of thirty (30) days. The initial seven (7) days were dedicated to distributing

the instruments to the respondents involved, while the remaining three (3) weeks were allocated for receiving the completed instruments for analysis.

3.1. Descriptive Analysis

Descriptive analysis was performed to identify the demographic characteristics of respondents, including gender, work experience, educational attainment, and other relevant variables. The analysis utilized frequency, percentage, and mean score values. The Statistical Package for the Social Sciences (SPSS) version 27 was employed for this purpose.

3.2. Inferential Analysis

Inferential analysis was used to answer the research questions and test the hypotheses. For this analysis, SmartPLS version 4.0 software was used to assess the measurement model, evaluate the structural model, compute the Variance Accounted For (VAF) for mediating effects, and examine the predictive accuracy of the overall model. Structural Equation Modeling (SEM) was applied to examine the theoretical relationships among the dimensions in the proposed model and simultaneous estimation of relationships between latent variables and their observed indicators. The process of SEM model development consists of two main stages: (1) development of the measurement model, and (2) development of the structural model.

3.2.1. Measurement Model

The measurement model was employed to test the observed variables derived from theory that represent the hypothesized latent constructs. It also explains the relationships between latent constructs and their respective indicators. This model further establishes the validity and reliability of the constructs. Table 1 presents the aspects and criteria applied in the measurement model.

Table 1. Aspects and Criteria for Measurement Model Evaluation

Aspect	Criteria
Internal Consistency Reliability	Composite Reliability (CR); Cronbach's Alpha
Convergent Validity	Outer Loading; Average Variance Extracted (AVE)
Discriminant Validity	Heterotrait-Monotrait Ratio of Correlations (HTMT)

Reliability refers to the consistency of scores obtained from an instrument, indicating that results remain stable and reproducible across different occasions (Creswell & Creswell, 2018). This study employed internal consistency reliability to assess the reliability of items, which measures the degree to which the items consistently represent the same construct. Composite Reliability (CR) is used to evaluate internal consistency. CR values range from 0 to 1, with higher values reflecting greater reliability. Specifically, CR values between 0.6 and 0.7 are acceptable for exploratory research (Nunnally & Bernstein, 1994) as shown at Table 2.

Table 2. Acceptable Range for Composite Reliability (CR)

CR Value	Criteria
> 0.78	Good
0.6 – 0.7	Acceptable for exploratory research
0.95	Accepted/rejected based on item review

Indicator reliability (outer loading) and Average Variance Extracted (AVE) were used to assess convergent validity, which evaluates the extent to which indicators of a construct correlate with each other. High variance explained by items within a construct demonstrates good construct validity.

According to Hair et al. (2019), as shown as at Table 3, an outer loading value of at least 0.708 is required since its squared value yields an indicator reliability of approximately 0.50—indicating that at least 50% of the variance in an indicator is explained by the corresponding construct.

Table 3. Criteria and Actions for Indicator Acceptance

Aspect	Value	Criteria and Action
Outer Loading	< 0.4	Item removed; effect on content validity considered
	0.4 – 0.70	Analyze impact of removal on CR and AVE: – If removal increases CR/AVE → remove item – If removal does not improve CR/AVE → retain item
	≥ 0.78	Item retained

Next, the Heterotrait-Monotrait (HTMT) ratio was applied to assess discriminant validity. HTMT measures the ratio between correlations within constructs and correlations across constructs. According to Kline (2016), discriminant validity is achieved when HTMT values are ≤ 0.85 . Furthermore, Henseler et al. (2016) stated that discriminant validity is confirmed if the confidence interval (CI) for HTMT does not include the value 1.0 for any construct.

3.2.2. Structural Model

The structural model was assessed after confirming that the measurement model was valid and reliable. This model represents the hypothesized relationships among latent constructs and was used to test the study's hypotheses. Once the measurement model was validated, the next step involved evaluating the model's predictive power and the significance of the hypothesized relationships. A systematic approach to evaluating the structural model in PLS-SEM includes collinearity assessment, path coefficient (β) significance testing, coefficient of determination (R^2), effect size (f^2) and predictive relevance (Q^2). The bootstrapping and PLSpredict procedures were conducted to obtain these values (Sarstedt et al. 2021). Table 4 shows the tests and criteria for Structural Model evaluation.

Table 4. Tests and Criteria for Structural Model Evaluation

No.	Test	Criteria and Acceptance Range
1.	Collinearity Assessment	Method: PLS Algorithm VIF < 5.0 = no multicollinearity VIF > 10 = potential multicollinearity issue (Hair et al., 2019)
2.	Path Coefficient (β) Significance	Method: Bootstrapping 1. $p < 0.01$, $t > 2.58$ (**), $t > 2.33$ (*) 2. $p < 0.05$, $t > 1.96$ (**), $t > 1.645$ (*) 3. $p < 0.10$, $t > 1.645$ (**), $t > 1.28$ (*) ** two tailed * one tailed
3.	Coefficient of Determination (R^2)	Range: 0–1; higher values indicate greater variance explained. (Chin, 1998) : $R^2 > 0.67$ = substantial, $R^2 > 0.30$ = moderate, $R^2 > 0.19$ = weak
4.	Effect Size (f^2)	Acceptance range according to (Chin, 1998; J. Cohen, 1988) : 0.02 = small; 0.15 = medium; 0.35 = large
5.	Predictive Relevance (Q^2)	Method: PLSpredict $Q^2 > 0$ indicates predictive relevance (Henseler et al., 2016)

Source: Adapted from Kamarudin Norazwa Yeop et al. (2021)

3.3. Sample of the study

The population of this study comprised senior vocational teachers (Heads of Programme) in Malaysian Vocational Colleges (N = 1,989) across 86 Vocational Colleges nationwide as of 2026 (BPLTV, 2026). These senior vocational teachers are responsible for monitoring teaching practices and ensuring alignment between curriculum, instruction, and assessment within their respective colleges.

Based on the sample size determination table proposed by Krejcie & Morgan (1970), a population exceeding 1,900 requires a minimum sample of 320 respondents. Because this study involves descriptive analyses to estimate means and standard deviations, the minimum sample size was increased by 20%, as recommended by Fraenkel et al. (2006). In addition, L. Cohen et al., (2017) suggested a minimum sample size of 322 for a population of approximately 2,000 at a 0.05 significance level (95% confidence level). Accordingly, the minimum required sample size for this study was set at 384 respondents.

To account for potential data loss during the PLS-SEM data screening and cleaning process, the study adopted an oversampling strategy, resulting in 504 questionnaires being distributed. All 504 questionnaires were returned. During data screening, 86 cases were excluded due to straight-lining patterns, and outliers. After the full PLS-SEM data-cleaning procedures were applied, 418 valid cases met all inclusion criteria and were retained for the final analysis. Table 5 show the sample selection and data screening for Heads of Programme.

Table 5. Sample Selection and Data Screening

Stage	Description	n
Population	Senior vocational teachers (Heads of Programme) in 86 Malaysian Vocational Colleges nationwide (TVET Division, 2026)	1,989
Sampling frame	Senior vocational teachers (Heads of Programme) in 86 Malaysian Vocational Colleges nationwide (TVET Division, 2026) + 20% as recommended by Fraenkel et al. (2006).	384
Invited to participate	Questionnaires distributed to eligible Heads of Programme across all strata	504
Responses received	Questionnaires returned	504
Excluded during data screening		86
	Straight-lining	9
	Multivariate outliers removed (Mahalanobis $p < .001$)	77
Final valid sample (N)	Cases meeting all PLS-SEM inclusion criteria and used in analysis	418

Because this study focuses on human behavioural responses, testing for common method bias is strongly recommended (Podsakoff et al., 2012). Accordingly, the Harman's single-factor test was conducted to assess the presence of common method variance. As shown in Table 6, the unrotated factor solution extracted 155 factors, with the largest single factor accounting for 47.19% of the total variance. Importantly, no single factor explained more than 50% of the variance, indicating the absence of a dominant general factor. Therefore, common method bias is not a serious concern in this study. While Harman's single-factor test has recognised limitations in detecting all forms of method variance, procedural remedies such as ensuring respondent anonymity were also implemented to minimise potential bias (Podsakoff et al., 2012).

Table 6. Summary of Factor Analysis for Common Method Bias

Factor	% Varian	Cumulative Variance Explained (%)
1	47.186	47.186

3.4 Validity and reliability of the instrument

The instrument has four sections: (1) Section A: Demographic Information (four items) Section B: B1. Technical Skills (20 items), B2. Industry 4.0 Generic Skills (55 items), B3. Career Adaptability Skills (23 items) and B4. TVET CA Practices (44 items). The instrument was initially developed based on an extensive literature review and validated for content validity by six (6) experts aligning with the number of experts in studies by (Mohd Matore & Khairani, 2020). These experts have over ten (10) years of working experience and from university lecturers holding a Doctor of Philosophy related to TVET. The minimum requirement for expertise in the field and research issue necessitated a duration of no less than three years, focusing on achieving consensus (Mohr & Shelton, 2017). This process ensures that the concept and information from the literature are accurately interpreted in the instrument (Drost, 2011).

The experts suggested adding several items, such as the ability to produce marketing literature and the use of online and contextualized learning activities. However, they also suggested removing an item such as the ability to deliver a speech to the audience. The Content Validity Ratio (CVR) is a statistical approach to evaluate content validity. The computation employs a formula that was developed by Lawshe (1975) and has gained popularity among academics for assessing the statistical significance of content validity (Azwan Masuwai et al., 2016; Nor Azizah Ahmad et al., 2019; Zamanzadeh et al., 2015). The CVR number is the mean level of appropriateness of the items derived from the comprehensive content validity panel. According to Mohd Matore & Khairani (2020), using the CVR value facilitates determining the retention or elimination of items, owing to the expert appraisal of content validity involved a three-point scale assigned to each instrument item, indicating the importance level. These points correspond to (1) Very Essential, (2) Less Essential, and (3) Not Essential. After collecting replies from all the panels, the number of issues the experts designated "Very Essential" was tallied (Ayre & Scally, 2014; Lawshe, 1975) and it has been determined that the crucial value of the CVR for a panel consisting of six experts is 1.

According to Lawshe (1975), there are two recommended approaches for updating or removing things that fall below the critical value of 1. These approaches involve either upgrading the items that fall below this value or eliminating them (Mohammed Afandi Zainal et al., 2020). Following the computation of the value for each CVR item, items exhibiting a value of 0.33 are excluded, items with a value of 0.67 necessitate enhancement, and items with a value of 1 are preserved. The number of items was reduced after the content validity assessment, and the number of items was subsequently adjusted and augmented per expert recommendations. The final items were going through evaluations conducted by experts on the item for each construct, leading the researchers to enhance the item following the advice provided by those experts.

After modification based on the experts' suggestions, the instrument was piloted to 114 TVET lecturers at Malaysian Polytechnic to ensure a high return rate and understanding of the instrument (Chua Yan Piaw, 2014). According to Hertzog (2008), it is recommended to employ a sample size ranging from 10 to 30 participants for instrument development during the pilot phase. The sample selection for this pilot study is based on a study sample in a context with almost the same characteristics as the actual study, namely among TVET trainers. An online questionnaire using the Google Form platform was also developed to facilitate quick feedback from the pilot study participants. Because the analysis of this study will use PLS-SEM, the data inspection method will follow the recommendations made by (Hair Jr. et al., 2021), such as exploratory analysis, which is Exploratory Factor Analysis (EFA), and calculate item reliability through Cronbach's alpha value.

The primary purpose of EFA is to summarize the data so that the relationships between the components are easily explained and understood (Yong & Pearce, 2013). EFA is also essential to test the factor structure hypothesis for the measurement set and find redundancy. EFA is a data exploration using principal component extraction with varimax rotation (variation maximization) on all items that measure each construct. Bartlett's Test of Sphericity was determined, and all constructs reached a significant value of $p < 0.05$.

In addition, the measurement of sample accuracy using Kaiser Meyer Olkin (KMO) is that all items must meet the set value above 0.5 (Kaiser, 1974). Based on Table 7, both results, which are KMO > 0.5 and Bartlett's Test of Sphericity, are significant for all of them, namely the construct of Technical Skills (TS) = (0.926), Industry 4.0 generic skills (IR4.0-GS) = (0.902), Career Adaptability Skills (CAS) = (0.954) and TVET CA Practices (CAP) = (0.911), and the data is suitable for proceeding to the following procedure.

Table 7. Finding of KMO and Bartlett's Test of Sphericity for Each Construct

Constructs/Type of Assessment		TS	IR4.0-GS	CAS	TVET CAP
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.926	0.902	0.954	0.911
Bartlett's Test of Sphericity	Approx. Chi-Square	1916.116	8743.728	2826.658	8392.237
	df	120	1891	190	1596
	Sig.	0.000	0.000	0.000	0.000

Next, the Total Variance Explained value for each construct is considered. The Total Variance Explained value at Table 8 shows that the number of items of each construct is appropriate to measure the construct. According to Pallant (2020), the overall value of the variant should exceed 50%. Table 8 shows that the total variance is over 50% for the construct of TS (65.40%), IR4.0-GS (56.04%), CAS (73.42%), and TVET CAP (58.08%).

Table 8. Total variance explained finding for each construct

Constructs	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% Variance	Cumulative %	Total	% Variance	Cumulative %
TS	10.465	65.409	65.409	10.465	65.409	65.409
IR4.0-GS	34.750	56.049	56.049	34.750	56.049	56.049
CAS	14.684	73.422	73.422	14.684	73.422	73.422
TVET CAP	33.109	58.086	58.086	33.109	58.086	58.086

Then, factor loading values in the component matrix for each construct are carried out. The higher the factor loading value, the higher the level of reliability. Cronbach's alpha values above 0.7 and factor loading values between 0.6 and 0.7 are accepted (Nunnally & Bernstein, 1994). Cronbach's alpha value for each variable and construct is above 0.8, and factor loading is above 0.6. This finding shows that the level of internal consistency is high, and the items are reliable.

4. Finding

4.1. Descriptive Analysis

This study uses SPSS 27 software to conduct a descriptive analysis divided into two analyses: descriptive analysis of respondents and analysis of the variables.

4.1.1. Descriptive Analysis of Respondents' Demographics

A descriptive demographic analysis of the respondents was conducted to obtain an overview of their demographic profiles. This section will report and discuss the profile of respondents who are senior teachers in Malaysian Vocational Colleges, including gender, state of service, descent, academic qualifications, length of service, period of teaching TVET subjects and TVET subjects taught.

The final analyzed data showed that 418 respondents from 5 states were involved in this study, representing their respective zones through stratified random sampling. A total of 24% of respondents

served in Perak, followed by 16% from Selangor, 19% of respondents served in Johor, and the rest served in other states. Most respondents, 77.0%, were female, while only 23.0% were male. In terms of ethnicity, the majority are Malay (95.9%). In addition, 4% of respondents have a Postgraduate Education Diploma, 92.1% have a bachelor's degree, and 6.7% have a master's degree. Based on length of service, 33.7% of respondents have served for 11-15 years, 27.8% have served for more than 20 years, 17.7% have served for 16-20 years, 11.7% have served for 6-10 years, and only 9.1% of respondents have served for less than six years.

4.1.2. Descriptive Analysis of the Variables

This section discusses and reports the descriptive analysis conducted on each variable in this study includes the mean and standard deviation. The instrument used is a five-point Likert scale that describes the level of agreement of the study participants to the questionnaire items, namely: 5 (strongly agree), 4 (agree), 3 (slightly agree), 2 (disagree), and 1 (strongly disagree). The variables involved are Technical Skills, Industry 4.0 Generic Skills, Career Adaptability Skills, and TVET CA Practices. Following that, the overall analysis of variables is also reported.

(a) Technical Skills Construct

A questionnaire consisting of 3 dimensions of Technical Skills, including method, process, and technique, was used to measure and evaluate Technical Skills. The findings of this study show that the teacher's agreement on the construct of Technical Skills is high based on the mean score (m)=4.10 and the standard deviation score (sd), which is 0.577 according to the dimension. The method dimension obtained a score of (m =4.07, sd =0.682), the process dimension received a score of (m =4.18, sd =0.590), and the technique dimension with a score of (m =4.06, sd =0.666).

(b) Industry 4.0 Generic Skill Constructs

The findings of this study show that the respondent's level of agreement on Industry 4.0 Generic Skill is at a moderate level based on the mean score value (m), which is 3.84, and 0.514 for the standard deviation score (sd) according to the dimension. The dimension of entrepreneurial skills obtained a score (m =3.68, sd =0.686), the dimension of interpersonal skills obtained a score (m =3.97, sd =0.605), the dimension of technology use skills obtained a score (m =3.77, sd =0.651), the dimension of teamwork skills received a score (m =3.97, sd =0.606), the dimension of leadership skills obtained a score (m =3.92, sd =0.609), the dimension of communication skills obtained a score (m =3.86, sd =0.583), the dimension of complex problem-solving skills received a score (m =3.64, sd =0.707), the dimension of technical competence and ICT knowledge obtained a score (m =3.85, sd =0.654), the dimension of communication skills and the ability to work with data obtained a score (m =3.74, sd =0.688) and the dimension of emotional intelligence and self-skills obtained a score (m =3.93, 0.591).

(c) Career Adaptability Skill Constructs

The findings of this study show that the respondent's level of agreement on Career Adaptability Skills is moderate based on the mean score of 3.76 and 0.630 for the standard deviation score according to the dimension. The dimension of concern obtained a score (m =3.67, sd =0.685), the dimension of curiosity obtained a score (m =3.71, sd =0.694), the dimension of control obtained a score (m =3.85, sd =0.661), and the dimension of confidence received a score (m =3.81, sd =0.666).

(d) TVET CA Practices Construct

The findings of this study show that the level of respondents' agreement with TEVT CA Practices is high based on the mean score of 4.14 and 0.476 for the standard deviation score according to the

dimension. The assessment concept dimension obtained a score ($m=4.10$, $sd=0.498$), the assessment approach dimension obtained a score ($m=4.20$, $sd=0.517$), and the Proficiency dimension obtained a score ($m=4.13$, $sd=0.544$).

4.2. Analysis of Initial Model

This study employed Partial Least Squares Structural Equation Modelling (PLS-SEM) using SmartPLS 4 to assess the reliability and validity of the measurement model prior to testing structural relationships. Given the multidimensional nature of the constructs, a two-stage higher-order construct approach (HOC) modelling was adopted, following the recommendations of Hair et al. (2019) and Henseler et al. (2016). Figure 2 illustrates the model for the first-order construct, which involves the analysis among the sub-constructs or dimensions of each construct.

In this model, Technical Skills (TS), Industry 4.0 Generic Skills (IR4.0 GS), Career Adaptability Skills (CAS), and Classroom Assessment Practices (CAP) were specified as second-order reflective–reflective constructs, each measured by several first-order dimensions. In the first stage, the latent variable scores of the first-order constructs were estimated. In the second stage, these scores were used as indicators to form the higher-order constructs.

Subsequently, after evaluating the first-order measurement model, several items/indicators were removed from the model due to indicator loading values of less than 0.7. The items/indicators removed were PRS23 and PUP1 under the assessment concept dimension. In addition, item KKKBD52 under the communication and data processing ability dimension was also excluded to achieve discriminant validity. Figure 3 illustrates the second-order measurement model involving all constructs in the study. The findings for both the first-order and second-order measurement models are discussed in the following section.

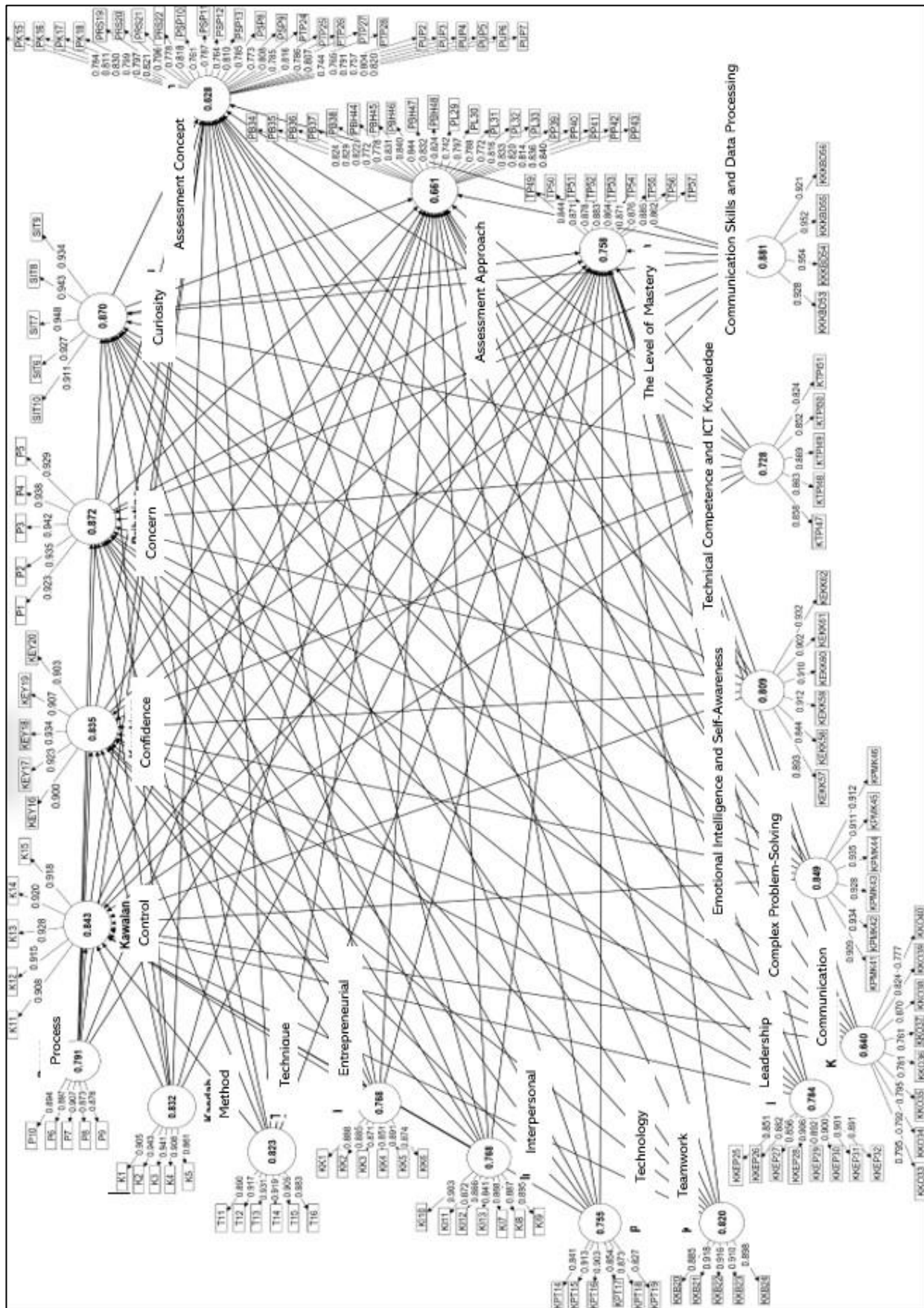


Figure 2. Model for the First-Order Construct

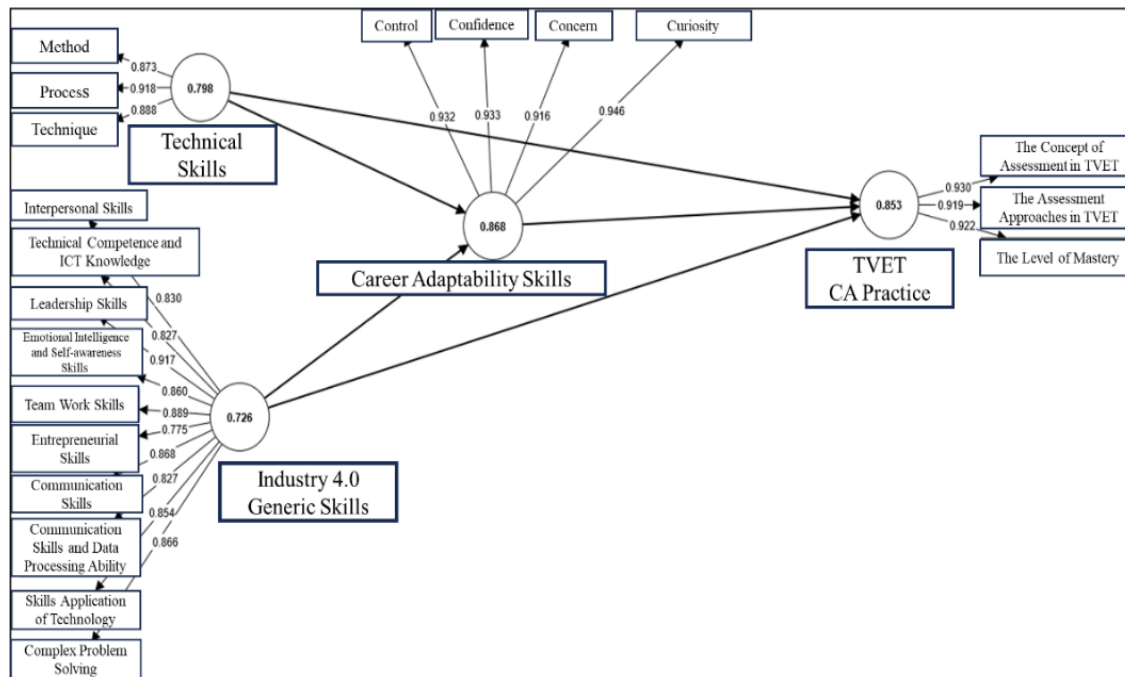


Figure 3. Second Order Measurement Model

4.2.1. Indicator Reliability

Indicator reliability was evaluated based on the outer loadings of each indicator on its respective construct. According to Hair et al. (2019), outer loadings should exceed 0.70, although values above 0.60 are acceptable in exploratory research. During the first-order measurement model assessment, three indicators—PRS2, PUP1, and KKKBD52—were found to have loadings below the acceptable threshold and were therefore removed from further analysis. After deletion, all remaining indicators exhibited loadings above 0.70, indicating satisfactory indicator reliability.

For the second-order constructs, all dimensions loaded strongly on their respective higher-order constructs. For example, Technical Skills was reflected by Method (0.911), Process (0.908), and Technique (0.922). Similarly, Generic Skills was represented by ten dimensions, including Communication (0.918), Leadership (0.904), Teamwork (0.899), and Problem Solving (0.897), all exceeding the recommended thresholds. Career Adaptability Skills was well represented by Concern (0.949), Curiosity (0.932), Control (0.920), and Confidence (0.946). Classroom Assessment Practices was reflected by Assessment Concept (0.950), Assessment Approach (0.945), and Assessment Mastery (0.957). These results confirm the adequacy of the higher-order measurement structure.

4.2.2. Internal Consistency Reliability

Internal consistency reliability was assessed using Cronbach's Alpha (CA) and Composite Reliability (CR). Values exceeding 0.70 indicate satisfactory reliability (Hair et al., 2019). All first-order constructs recorded CA and CR values above the threshold, indicating strong internal consistency. At the second-order level, Technical Skills (CR = 0.931), Generic Skills (CR = 0.948), Career Adaptability Skills (CR = 0.961), and Classroom Assessment Practices (CR = 0.963) all demonstrated excellent reliability. These results confirm that the constructs consistently measure their intended latent variables.

4.2.3. Convergent Validity

Convergent validity was evaluated using the Average Variance Extracted (AVE). An AVE value of 0.50 or higher indicates that the construct explains more than half of the variance of its indicators (Bagozzi & Yi, 2012). All first-order and second-order constructs achieved AVE values above 0.50. At the higher-order level, Technical Skills (AVE = 0.818), Generic Skills (AVE = 0.645), Career Adaptability Skills (AVE = 0.860), and Classroom Assessment Practices (AVE = 0.898) demonstrated strong convergent validity.

4.2.4. Discriminant Validity

Discriminant validity was assessed using the Heterotrait–Monotrait Ratio (HTMT). According to Henseler et al. (2016), HTMT values below 0.90 (or below 1.00 for conceptually related higher-order constructs) indicate acceptable discriminant validity. The HTMT values among the higher-order constructs ranged from 0.626 to 0.920 as shown at Table 9. Although the HTMT value between IR 4.0 Generic Skills and Career Adaptability Skills reached 0.920, this is considered acceptable given their conceptual proximity and higher-order nature. Importantly, no HTMT value exceeded 1.00, confirming that all constructs are empirically distinct.

Table 9. Heterotrait-Monotrait Ratio (HTMT) values of first-order constructs

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2	0.645																			
3	0.749	0.601																		
4	0.609	0.457	0.711																	
5	0.818	0.633	0.826	0.734																
6	0.733	0.573	0.806	0.722	0.812															
7	0.798	0.590	0.764	0.727	0.883	0.774														
8	0.745	0.600	0.673	0.557	0.741	0.671	0.695													
9	0.700	0.551	0.884	0.713	0.818	0.811	0.775	0.634												
10	0.700	0.543	0.780	0.806	0.828	0.769	0.784	0.648	0.756											
11	0.581	0.421	0.694	0.877	0.725	0.699	0.691	0.598	0.698	0.742										
12	0.573	0.588	0.692	0.573	0.636	0.607	0.620	0.480	0.660	0.573	0.500									
13	0.482	0.477	0.565	0.509	0.560	0.520	0.526	0.360	0.539	0.525	0.434	0.812								
14	0.746	0.559	0.722	0.747	0.786	0.705	0.804	0.690	0.704	0.750	0.708	0.580	0.517							
15	0.663	0.511	0.704	0.748	0.799	0.750	0.775	0.655	0.765	0.777	0.780	0.512	0.428	0.739						
16	0.666	0.515	0.815	0.694	0.767	0.754	0.717	0.630	0.832	0.717	0.714	0.539	0.462	0.715	0.762					
17	0.667	0.747	0.619	0.514	0.647	0.593	0.621	0.600	0.570	0.592	0.470	0.565	0.492	0.608	0.525	0.560				
18	0.711	0.576	0.878	0.715	0.789	0.769	0.747	0.641	0.874	0.724	0.721	0.638	0.526	0.725	0.783	0.882	0.556			
19	0.591	0.556	0.679	0.544	0.608	0.612	0.593	0.451	0.631	0.542	0.465	0.798	0.802	0.549	0.501	0.560	0.564	0.618		
20	0.689	0.662	0.610	0.521	0.671	0.585	0.605	0.639	0.578	0.617	0.465	0.512	0.439	0.682	0.558	0.589	0.795	0.572	0.496	

Note: 1. Interpersonal Skills 2. Method 3. Control 4. Technical Competence and ICT Knowledge 5. Leadership Skills 6. Emotional Intelligence and Self-Awareness Skills 7. Teamwork Skills 8. Entrepreneurial Skills 9. Confidence 10. Communication Skills 11. Communication Skills and Data Processing Ability 12. Concept of Assessment 13. Assessment Approaches 14. Skills in the Application of Technology 15. Complex Problem-Solving Skills 16. Concern 17. Process 18. Curiosity 19. Level of Mastery 20. Technique

Table 10 presents the discriminant validity values for the second-order constructs as measured by the Heterotrait-Monotrait Ratio (HTMT). The HTMT value assesses the mean correlation of indicators across constructs. According to Henseler et al. (2016), discriminant validity is established when the HTMT value is less than 0.90. However, for higher-order models, HTMT values approaching 1 are still acceptable (HTMT < 1). Based on Table 10, all HTMT values for the second-order constructs are below 1. Therefore, the discriminant validity criterion has been satisfactorily met.

Table 10. Heterotrait-Monotrait Ratio (HTMT) value for second-order construct

	TVET CA Practices	IR 4.0-GS	CAS	TS
TVET CA Practices				
IR 4.0-GS	0.693			
CAS	0.713	0.920		
TS	0.679	0.790	0.723	

4.3. Analysis of Measurement Model

According to Hair et al., (2014), it can generally be said that the evaluation of the measurement model can be explained based on the assessment of the reliability and validity of items/constructs and constructs. Table 11 shows the findings of the reliability and validity assessment. Figure 4 provides a summary of the initial measurement model for second order constructs.

Table 11. The Assessment of the Initial Measurement Model

	Item	Outer loading	Alpha Cronbach	CR
Technical Skills	Method	0.873	0.949	0.961
	Process	0.918	0.934	0.950
	Technique	0.888	0.957	0.965
Industry 4.0 Generic Skills	Entrepreneurial Skills	0.775	0.940	0.952
	Interpersonal Skills	0.830	0.949	0.959
	Skills Application of Technology	0.854	0.935	0.949
	Teamwork Skills	0.889	0.945	0.958
	Leadership Skills	0.917	0.960	0.967
	Communication Skills	0.868	0.919	0.934
	Complex Problem Solving	0.866	0.964	0.971
	Technical Competence and ICT Knowledge	0.827	0.907	0.930
	Communication Skills and Data Processing Ability	0.827	0.943	0.967
	Emotional Intelligence and Self-awareness Skills.	0.860	0.952	0.962
Career Adaptability Skills	Concern	0.916	0.963	0.971
	Curiosity	0.946	0.963	0.971
	Control	0.932	0.953	0.964
	Confidence	0.933	0.950	0.962
TVET CA	The Concept of Assessment in TVET	0.930	0.977	0.978
Practices	The Assessment Approaches in TVET	0.919	0.973	0.975
	The Level of Mastery	0.922	0.960	0.966

4.3.1. Analysis of the Significance Affect between the Variables

RQ1: To what extent do teachers perceive students' Technical Skills as influencing their Career Adaptability Skills?

The finding shows the path coefficient value of the structural model to give an overview of the hypothetical relationship between the constructs in the structural model. Hair et al. (2014) suggest a significant level that can be accepted is at least at a significant level of 10%. Based on the findings shows that Technical Skills significantly positively affect Career Adaptability skills ($p < 0.10$). The path value of the coefficient ($\beta = 0.050$, $t = 1.339$) explains that each unit increase in Technical Skills will increase by about 0.050 Career Adaptability Skills. This shows that Technical Skills have a positive and significant relationship and subsequently influence Career Adaptability Skills in the TVET CA Practices. Respondents believe that students with a high level of Technical Skills will have a high level of Career Adaptability Skills, which are very suitable to be applied in the TVET CA practices.

RQ2: To what extent do teachers perceive students' Industry 4.0 Generic Skills as influencing their Career Adaptability Skills?

The finding shows Industry 4.0 Generic Skills significantly positively affect Career Adaptability skills ($p < 0.05$). The path value of the coefficient ($\beta = 0.843$, $t = 29.115$) explains that each unit of increase in Industry 4.0 Generic Skills will increase 0.843 Career Adaptability Skills. This shows that Industry 4.0

Generic Skills have a positive and significant relationship and subsequently affect Career Adaptability Skills in the TVET CA Practices. The results show that students with a high level of Industry 4.0 Generic Skills will have a high level of Career Adaptability Skills, which are suitable to be applied in the TVET CA Practices.

RQ3: How do teachers perceive the influence of students' Career Adaptability Skills on TVET CA Practices?

The finding shows Career Adaptability Skills significantly positively affect the TVET CA Practices ($p < 0.05$). The path value of the coefficient ($\beta = 0.389$, $t = 5.007$) explains that each unit of increase in career adaptability skills will increase 0.389 of TVET CA Practices. The finding shows that Career Adaptability Skills are positively and significantly related to the TVET CA Practices.

RQ4: How do teachers perceive the influence of students' Technical Skills on TVET CA Practices?

The path value of the coefficient ($\beta = 0.268$, $t = 5.039$) explains that each unit of increase in Technical Skills will increase 0.268 TVET CA Practices. It showed that positive relationship Technical Skills influence the TVET CA Practices significantly.

RQ5: How do teachers perceive the influence of students' Industry 4.0 Generic Skills on TVET CA Practices?

The finding shows that Technical Skills significantly positively affect the TVET CA Practices ($p < 0.05$). Industry 4.0 Generic Skills marginally significant, indicating a relatively weak effect the TVET CA practices ($p < 0.10$). The coefficient path value ($\beta = 0.119$, $t = 1.467$) explains that each unit increase in Industry 4.0 Generic Skills will increase 0.119 TVET CA Practices. The finding shows that Industry 4.0 Generic Skills have a relatively weak relationship to TVET CA Practices. Table 12 shows significance effect of each variable for the structural model.

Table 12. Significance Effect of Each Variable for the Structural Model

		Path Coefficient	t value	p-value	LCL	UCL
H_{a1}	Technical Skills → Career Adaptability Skills	0.050	1.339	0.090*	0.003	0.098
H_{a2}	Industry 4.0 Generic Skills → Career Adaptability Skills	0.843	29.115	0.000***	0.795	0.891
H_{a3}	Career Adaptability Skills → TVET CA Practices	0.389	5.007	0.000***	0.261	0.515
H_{a4}	Technical Skills → TVET CA Practices	0.268	5.039	0.000***	0.182	0.356
H_{a5}	Industry 4.0 Generic Skills → TVET CA Practices	0.119	1.467	0.071*	0.016	0.228

Note: ***Significant at 1%, * Significant at 10%, LCL- lower confidence level, UCL- upper confidence level

4.3.2. Analysis of Structural Model based on the Value of Predictive Accuracy (R²), Predictive Relevance (Q²) and Effect Size (f²)

The findings obtained are by making a summary of the relationships that exist between all the indicators and constructs involved. Once the model is developed, the next test will focus on the accuracy of prediction (R²), the actual effect (f²) and the relevant effect (Q²) of the model. Coefficient of determination (R²) is a coefficient of determination to measure the accuracy of model predictions and is calculated as a squared correlation between the actual value and the predicted value of the dependent variable. The R² explains the variation (change) in the dependent (endogenous) variable that is explained by the amount of change in the independent (exogenous) variable. In other words, it represents the sum of variations that are related to each other by systematically sharing the same pattern. The R² value is between 0 and 1 and the higher the R² value, the higher the expected accuracy level. According to Chin (1998), R² > 0.67 is strong, R² > 0.3 is moderate, and R² > 0.19 is weak. Table 12 shows the value of the R² for TVET CA Practices and Career Adaptability Skills. Based on the value

of R², 50.1% of the total variation in TVET CA Practices is explained by Technical Skills, Industry 4.0 Generic Skills and Career Adaptability Skills, while for the total variation of Career Adaptability Skills is as much as 77.4%.

In addition to evaluating R² for all dependent variables, changes in R² values when certain independent variables are excluded from the model can be used to evaluate whether the predictor construct has a substantive (real) effect on the dependent variable. This method is called effect size (f^2). Table 13 shows finding of effect size (f^2). The finding showed that 77.4% of the total variance of Career Adaptability Skills is explained by Technical Skills and Industry 4.0 Generic Skills. Based on the effect size value, there is no effect between Industry 4.0 Generic Skills on TVET CA Practices, while there is a small relationship between Career Adaptability Skills and Technical Skills on TVET CA Practices. Next, by using the PLSpredict procedure, the predictive relevance is obtained to determine the strengthening of the model (Q²). According to Henseler et al., (2009, 2016) a Q² value less than 0 means that the model has less predictive relevance. Based on the Q² value, this model is proven to have predictive relevance. Therefore, the Figure 4 is Structural Model based on the value of predictive accuracy (R²), predictive relevance (Q²) and effect size (f^2).

Table 13. Effect Size (f^2)

Constructs	TVET CA Practices	
	f^2	Size
Industry 4.0 Generic Skills	0.005	No effect
Career Adaptability Skills	0.069	Small
Technical Skills	0.068	Small

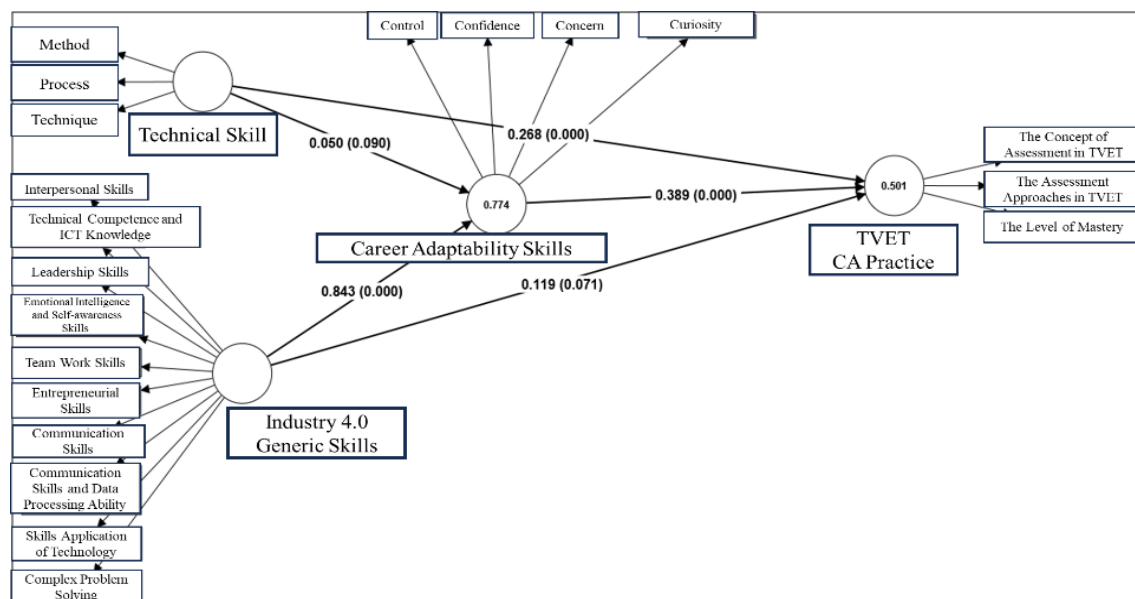


Figure 4. Structural Model

RQ6: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Technical Skills and CA Practices?

RQ7: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Industry 4.0 Generic Skills and CA Practices?

The structural model path coefficient in Table 14 shows that Career Adaptability Skills are not significant mediators in the relationship between Technical Skills and TVET CA Practices [$(p > 0.05)$, $\beta = 0.019$, $t = 1.275$]. The findings show that Career Adaptability Skills do not play an essential role in

increasing the effect of Industry 4.0 Generic Skills on TVET CA Practices. Therefore, this alternative hypothesis statement is not supported. The findings of the study prove that Career Adaptability Skills are a significant mediator in the relationship between Industry 4.0 Generic Skills and TVET CA Practices [($p < 0.01$), $\beta = 0.328$, $t = 4.934$]. Based on Table 14, the increase in the value of the coefficient path, $\beta = 0.328$, also shows that Career Adaptability Skills are essential to increase the effect of Industry 4.0 Generic Skills in the TVET CA Practices. Respondents define that Industry 4.0 Generic Skills and Career Adaptability Skills influence each other, and applying these skills is essential in the TVET CA Practices.

Table 14. Mediation Analysis Result

		Path Coefficient	t value	p-value	LCL
H _{a6}	Technical Skills → Career Adaptability Skills → TVET CA Practices	0.019	1.275	0.101	-0.003
H _{a7}	Industry 4.0 Generic Skills → Career Adaptability Skills → TVET CA Practices	0.328	4.934	0.000***	0.222

Note: ***Significant at 1% level

5. Discussion

5.1. Technical Skills and Career Adaptability Skills (Ha1)

RQ1: To what extent do teachers perceive students' Technical Skills as influencing their Career Adaptability Skills?

The findings indicate that the relationship between Technical Skills and Career Adaptability Skills is marginally significant, suggesting a relatively weak association between these constructs. From the perspective of TVET teachers, students who demonstrate stronger technical competencies may exhibit slightly higher levels of adaptability; however, technical skills alone may not be sufficient to fully explain variations in career adaptability. This finding suggests that while technical competencies remain important in vocational education, they may not directly translate into adaptability-related attributes such as confidence, curiosity and flexibility. This is consistent with previous studies which emphasise that adaptability is influenced by a broader set of competencies beyond technical expertise. Therefore, the role of technical skills in shaping career adaptability should be interpreted with caution.

In TVET, developing Technical Skills is essential to prepare students to enter the world of work. However, Technical Skills alone are not enough to be successful in their career choices. Career Adaptability Skills such as concern, curiosity, control, and confidence are also needed to face changes and career development (Savickas & Porfeli, 2012; Zool Hilmi Mohamed Ashari, 2016). The findings of this study differ from those conducted by (Ebenehi et al., 2016); that is, students with Technical and Vocational skills do not significantly affect career adaptability skills with a finding of $\beta = .02$, $p = .653$. However, Technical Skills in the VET sector can help students acquire the career adaptability skills needed to face the needs of the ever-evolving world of work (de Guzman & Choi, 2013).

TVET CA Practices plays a vital role in forming and developing students' Career adaptability skills. This is because the TVET CA practices involve assessing the students' Technical Skills. Pupils are tested based on their performance in tasks and practical work related to the vocational field studied. This assessment process helps determine the students' Technical Skill levels and provides feedback for improvements that stakeholders can implement, especially trainers and the school. The TVET CA Practices, according to Lahn & Nore (2019) and Rauner et al. (2013) also provides an opportunity to cultivate students' Career Adaptability Skills through several methods and findings, such as continuous learning and self-management.

5.2. Industry 4.0 Generic Skills and Career Adaptability Skills (Ha2)

RQ2: To what extent do teachers perceive students' Industry 4.0 Generic Skills as influencing their Career Adaptability Skills?

The results show that Industry 4.0 Generic Skills have a strong and significant relationship with Career Adaptability Skills, indicating that this is the most influential predictor in the model. Based on teachers' observations, students who demonstrate higher levels of critical thinking, problem-solving and adaptability are more likely to exhibit stronger career adaptability. This finding highlights the importance of transferable competencies in preparing students for dynamic and uncertain work environments. Unlike technical skills, Industry 4.0 generic skills appear to play a central role in enabling students to manage career transitions and respond to evolving job demands. This supports existing literature which emphasises the importance of soft and generic skills in the context of Industry 4.0.

In line with the findings of the multivariate test conducted by de Guzman & Choi (2013), there is a positive relationship between Career Adaptability Skills and Industry 4.0 Generic Skills through communication skills with values ($r=.48$, $pb.001$) and problem-solving skills with values ($r=.44$, $pb.001$). This study also shows that Career Adaptability Skills correlate significantly with teamwork ($r=.67$, $pb.001$). The findings of de Guzman & Choi (2013) showed that the four elements of Career Adaptability Skills positively correlate with Industry 4.0 Generic Skills such as communication, problem-solving, and teamwork. The study by Ebere & Onuoha (2022) showed that there is a significant and positive relationship between leadership skills and Career Adaptability Skills with values ($p=0.000 < 0.05$ $r=0.700$). This study also found that problem-solving skills have a significant and positive relationship with value ($p=0.000 < 0.05$) and $r=0.778$).

This finding shows that Industry 4.0 Generic Skills are critical in forming Career Adaptability Skills in the TVET assessment curriculum, especially for the TVET CA practices. Industry 4.0 Generic Skills and Career Adaptability Skills enable students to adapt to changing job demands and improve their efficiency at work when venturing into a career field. Students with high Industry 4.0 Generic Skills are seen as more capable of adapting to job demands, solving problems, and collaborating well with friends. Following the statement Acerbi et al., (2022), the development of Industry 4.0 Generic Skills in today's world of education is increasing towards digitization, networking, significant life changes, and even changes in the world of work today also shift the paradigm in education.

5.3. Career Adaptability Skills and Classroom Assessment Practices (Ha3)

RQ3: How do teachers perceive the influence of students' Career Adaptability Skills on TVET CA Practices?

The findings indicate that Career Adaptability Skills have a significant positive relationship with Classroom Assessment Practices. From teachers' perspectives, students who demonstrate higher levels of adaptability are more likely to perform effectively in assessment activities. This suggests that adaptable students are better able to engage with assessment tasks, respond to feedback and demonstrate their competencies in various learning contexts. As classroom assessment in TVET involves performance-based evaluation, students' ability to adapt to different tasks and expectations plays an important role in determining their performance outcomes.

In TVET, CA Practices is essential in developing and measuring Career Adaptability Skills. For example, the Human Resources Development (HRD) model by McLagan (1989) is an evaluation framework that measures human resource competence, especially skills involving Technical Skills and Industry 4.0 Generic Skills. Then, the findings from the evaluation help provide appropriate development programs for career advancement (Muhd Khaizer Omar et al., 2021). The practices of TVET Classroom Assessment can support the development of Career Adaptability Skills in the field of TVET as suggested by the Competence Assessment and Development Model (COMET) Rauner et al.

(2013), i.e., skills and competencies should be assessed with accurate assessment methods to support the development of human resource competencies.

CA Practices can be implemented through (1) An assessment concept allowing students to use their Career Adaptability Skills in the real world. For example, outcome-based assessment and observation-based approaches where students are required to solve real-world problems can help them develop their problem-solving skills, teamwork, and adaptability. (2) Formative assessment provides continuous feedback on student progress, enabling them to identify areas for improvement and make appropriate judgments. Students can improve their abilities and confidence by receiving feedback on their Career Adaptability Skills. (3) Self-assessment: Self-assessment and reflection can help students develop self-awareness and identify their strengths and weaknesses. By considering their skills, interests, and values, pupils can make important decisions about their career paths and develop strategies to adapt to an ever-changing learning and career environment. In addition, career assessment tests can also be conducted to help students explore different career options and identify their interests and strengths. These tests can include personality tests, interest inventory tests, and student skills assessments.

5.4. Technical Skills and Classroom Assessment Practices (Ha4)

RQ4: How do teachers perceive the influence of students' Technical Skills on TVET CA Practices?

The results show that Technical Skills have a significant positive relationship with Classroom Assessment Practices. This finding is expected, as technical competencies are directly evaluated in TVET assessment activities. Teachers assess students' ability to perform practical tasks, apply technical knowledge and demonstrate job-related competencies. Therefore, students who exhibit stronger technical skills are more likely to perform well in classroom assessment practices. This finding reinforces the importance of competency-based assessment approaches in TVET education.

Technical skills are essential applied in TVET CA Practices, specifically for TVET subjects, because they allow students to master the skills needed to carry out tasks and responsibilities in the field they are studying. TVET aims to equip students with the Technical Skills required to become a potential workforce (Fah, 2018). TVET CA Practices are an essential aspect of the TVET curriculum because the right CA Practices can help trainers assess student progress and ensure that they have acquired the necessary Technical Skills per the TVET field (Ridzwan Che Rus et al., 2017). Here are some TVET CA Practices that can apply Technical Skills in the field of TVET, as suggested by Sutarto & Jaedun (2018): (1) outcome-based assessment, which measures students' ability to use their Technical Skills in practical situations. This assessment can include building structures, solving technical problems, or handling equipment and materials. Assessment in practical training or 'hands-on' gives students practical experience in using their Technical Skills (Knight & Sweet, 2010). This assessment can include tasks such as cooking, sewing, assembling electronic equipment, welding two pieces of metal together, or doing laboratory experiments related to TVET, and (2) the concept of assessment through formative and summative assessment can provide continuous feedback to students and allow them to identify areas for improvement according to their potential.

By receiving feedback on students' Technical Skills, they can improve their abilities and build their confidence. CA Practices is essential in supporting the development of Technical Skills in the TVET curriculum (McSweeney, 2014). The variety of assessment concepts in the TVET field, such as competency-based assessment, performance-based assessment, practical assessment, and formative assessment, are effective methods to measure and support the development of Technical Skills in the TVET field (Niar Rosmaliana Rusalam et al., 2019). However, classroom assessment that adapts formative assessment and summative assessment is also an assessment that can measure the domain of student learning outcomes from the aspects of knowledge, skills, and attitudes as well as based on the actual situation of the work environment (Sephokgole & Makgato, 2019). By adopting TVET CA Practices, trainers can ensure that their students are prepared to learn and subsequently succeed in the job sector they are engaged in.

5.5. Industry 4.0 Generic Skills and Classroom Assessment Practices (Ha5)

RQ5: How do teachers perceive the influence of students' Industry 4.0 Generic Skills on TVET CA Practices?

The relationship between Industry 4.0 Generic Skills and Classroom Assessment Practices was found to be marginally significant, indicating a relatively weak effect. While students who demonstrate stronger generic skills may show better engagement in assessment activities, the direct influence of these skills on assessment outcomes appears to be limited. This suggests that Industry 4.0 generic skills may not be directly reflected in assessment performance unless they are explicitly integrated into assessment design. Therefore, the role of generic skills in classroom assessment should be further strengthened through more holistic and integrated assessment approaches.

Assessment practices in implementing the TVET program aim to equip students with the skills and knowledge needed to succeed in the world of work. Technical skills are a critical requirement in TVET, but there is also a growing need for Industry 4.0 Generic Skills. These elements of Industry 4.0 Generic Skills are seen as transferable requirements across different industries and job types, such as communication, problem-solving, teamwork, and leadership skills. CA practices are essential in developing and measuring these Industry 4.0 Generic Skills, especially in the TVET curriculum.

The concept of assessment in the form of summative assessment allows students to use their skills and knowledge in a real-world context (McSweeney, 2014). For example, implementing assessments in the form of group projects and presenting project results where students are required to work together to solve real problems can help them develop teamwork, interpersonal, problem-solving, and leadership skills and further develop communication skills. According to Sandal et al. (2014), the practical aspect of the assessment in the field differs from the theoretical assessment students learn in the classroom. Their exposure to practical activities helps build communication skills, responsibility, and proper work ethics. In addition, implementing assessment through oral assessment or demonstration can help students develop communication skills. This assessment can also help students build confidence, measure their abilities, and increase their added value.

5.6. Mediating Role of Career Adaptability Skills (Ha6 & Ha7)

RQ6: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Technical Skills and CA Practices?

RQ7: According to teachers' perceptions, do Career Adaptability Skills mediate the relationship between Industry 4.0 Generic Skills and CA Practices?

The mediation analysis indicates that Career Adaptability Skills play a mediating role in the relationships between competency domains and Classroom Assessment Practices. Specifically, Career Adaptability Skills partially mediate the relationship between Industry 4.0 Generic Skills and Classroom Assessment Practices. This finding suggests that Industry 4.0 generic skills contribute to assessment outcomes indirectly by enhancing students' adaptability. Students who possess strong generic competencies are more likely to develop adaptability, which in turn supports their performance in assessment activities.

However, the mediating effect between Technical Skills and Classroom Assessment Practices is relatively weak, indicating that technical skills may influence assessment outcomes more directly rather than through adaptability. Overall, these findings highlight the importance of considering both direct and indirect pathways in understanding competency development in TVET.

The findings of this study show that Career Adaptability Skills mediate the relationship between generic skill constructs and TVET CA Practices. However, Career Adaptability Skills are not intermediate

between Technical Skills, and Career Adaptability Skills refer to an individual's ability to adapt to the ever-changing needs of the job market. Industry 4.0 Generic Skills constructs can help students acquire Career Adaptability Skills. The study of Cortellazzo et al. (2020) proved a positive relationship between Career Adaptability Skills and Industry 4.0 Generic Skills with value ($\beta = .46, p < .01$). The findings of this study also demonstrate that students who have Industry 4.0 Generic Skills and high Career Adaptability Skills have the potential to get suitable jobs and succeed in entering the field of employment. Constructs of Industry 4.0 Generic Skills such as communication, teamwork, and complex problem-solving skills help students improve the skills needed in jobs involving other people and solving complex problems. Recognizing childhood as the beginning of the development of vocational skills and the basis for Career Adaptability, Hartung et al. (2008) assert that Career Adaptability Skills can be developed as early as childhood.

The development of technology and society is always linked to continuous changes in the environment and the role and structure of work. An individual's identity is usually formed from childhood and will continue to develop in their community and environment. Therefore, in this situation, the school environment can prepare students to face the world of work (Savickas & Porfeli, 2012). CA Practices for TVET subjects can also improve students' Career Adaptability Skills. Classroom assessment in the field of TVET allows trainers to assess a student's knowledge and skills involving authentic teaching and learning situations. In this process, students will experience challenges, face problems, and learn to make decisions by considering various factors. Therefore, through applying Industry 4.0 Generic Skill constructs in classroom assessment practices, students successfully acquire the Career Adaptability Skills needed to succeed in the work they are engaged in. Career Adaptability Skills mediate this relationship, as they play an essential role in linking Industry 4.0 Generic Skills and helping to equip students to meet the challenges of learning and thus improve their competence in the ever-changing real world of work.

6. Conclusion

This study examines the relationships between technical skills, Industry 4.0 generic skills, career adaptability skills and classroom assessment practices in Technical and Vocational Education and Training (TVET) based on teachers' perceptions of students' competencies. The findings provide empirical evidence on how different competency domains interact within TVET learning environments and influence classroom assessment practices.

The results indicate that both technical skills and career adaptability skills have significant positive relationships with classroom assessment practices, highlighting their direct importance in competency-based evaluation. In contrast, Industry 4.0 generic skills demonstrate a relatively weaker, marginally significant relationship with classroom assessment practices, suggesting that these competencies may not be fully captured in current assessment practices. In addition, Industry 4.0 generic skills show a strong and significant relationship with career adaptability skills, while technical skills exhibit a comparatively weaker association. These findings suggest that adaptability-related competencies are more strongly shaped by transferable and generic skills than by technical competencies alone.

Furthermore, the findings reveal that career adaptability skills partially mediate the relationship between Industry 4.0 generic skills and classroom assessment practices, indicating an indirect pathway through which generic competencies influence assessment outcomes. However, the mediating effect of career adaptability skills in the relationship between technical skills and classroom assessment practices is relatively weak, suggesting that technical competencies are more directly assessed within TVET classroom settings.

From a theoretical perspective, this study contributes to the literature by integrating multiple competency constructs into a unified structural model and highlighting the mediating role of career adaptability skills in competency development. The findings extend existing knowledge by

demonstrating that competency development in TVET is multidimensional and requires the integration of both technical and generic skill domains.

From a practical perspective, the findings suggest that TVET institutions should strengthen the integration of Industry 4.0 generic skills within teaching, learning and classroom assessment practices. While technical skills remain essential, current assessment approaches may need to be enhanced to better capture transferable competencies such as problem-solving, adaptability and communication. In addition, greater emphasis should be placed on developing career adaptability skills, as these competencies play a key role in supporting students' performance in assessment and their readiness for the workforce.

This study has several limitations. First, the findings are based on teachers' perceptions of students' competencies rather than direct measurement of students' abilities. Future studies may incorporate multiple data sources, including student self-assessment and industry evaluation, to provide a more comprehensive perspective. Second, the study is limited to a specific TVET context, which may affect the generalisability of the findings. Future research may also consider employing qualitative approaches to gain deeper insights into classroom assessment practices and competency development in TVET.

In conclusion, this study underscores the importance of aligning competency development with classroom assessment practices in TVET. By enhancing the integration of technical skills, Industry 4.0 generic skills and career adaptability skills, TVET institutions can better support the development of competent and adaptable graduates who are prepared to meet the evolving demands of the workforce.

Acknowledgement

Not applicable.

Declaration

The authors declare that minimal use of artificial intelligence tools was limited to language editing only. All scholarly content and interpretations are the authors' own. Ethical approval for this study was obtained from Technical Vocational Education and Training Division, Ministry of Education with the reference number KPM.600-3/2/3 Jld.2 (84)

References

- Acerbi, F., Rossi, M., & Terzi, S. (2022). Identifying and Assessing the Required I4.0 Skills for Manufacturing Companies' Workforce. *Frontier in Manufacturing Technology*, 2(July), 1–19. <https://doi.org/10.3389/fmtec.2022.921445>
- Adam, Asyraf, Farisha, Aisyah, & Ainun. (2019). Exam-oriented in Malaysia's Education System. *E-Journal of Media & Society*, 3(1), 1–19. www.ejoms.com
- Affero Ismail, Wan Azrul Shafuan Wan Hassan, Muhd Fazrul Helmi Ahmad, Affan Z, & Harun, M. I. (2020). Students' Readiness in Facing Industrial Revolution 4.0 Among Students of Technical Teacher's Education. *International Journal of Scientific & Technology Research*, 9(8), 300–305. www.ijstr.org
- Ahmad Zaidi Husin. (2019). *Model Amalan pentaksiran Formatif dalam Kalangan Guru-Guru Kemahiran Hidup Bersepadu* [Universiti Utara Malaysia]. <http://myto.upm.edu.my/find/Record/myuum-etd.8109>
- Aizuddin Saari, Mohamad Sattar Rasul, Ruhizan Mohamad Yasin, Rose Amnah Abd Rauf, Zool Hilmi Mohamed Ashari, & Diaz Pranita. (2021). Skills Sets for Workforce in the 4th Industrial Revolution: Expectation From Authorities and Industrial Players. *Journal of Technical Education and Training*, 13(2), 1–9. <https://doi.org/10.30880/jtet.2021.13.02.001>
- Ana, A., Yulia, C., Jubaedah, Y., Muktiarni, M., Dwiyantri, V., & Maosul, A. (2020). Assessment of Student Competence Using Electronic Rubric. *Journal of Engineering Science and Technology*, 15(6), 3559–3570.
- Asnidatul Adilah Ismail, & Razali Hassan. (2019). Technical Competencies in Digital Technology Towards Industrial Revolution 4.0. *Journal of Technical Education and Training*, 11(3), 55–62. <https://doi.org/10.30880/jtet.2019.11.03.008>
- Ayre, C., & Scally, A. J. (2014). Critical Values for Lawshe's Content Validity ratio: Revisiting the Original Methods of Calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86. <https://doi.org/10.1177/0748175613513808>
- Azmanirah Ab Rahman, Nurfirdawati Muhammad Hanafi, Yusmarwati Yusof, Ibrahim Mukhtar Marina, Halizah Awang, & Anizam Mohamed Yusof. (2020). The Effect of Rubric on Rater's Severity and Bias in TVET Laboratory Practice Assessment: Analysis Using Many-Facet Rasch Measurement. *Journal of Technical Education and Training*, 12(1), 57–67. <https://doi.org/10.30880/jtet.2020.12.01.006>
- Azwani Masuwai, Nor'ain Mohd Tajudin, & Noor Shah Saad. (2016). Evaluating the Face and Content Validity of a Teaching and Learning Guiding Principles Instrument (TLGPI): a Perspective Study of Malaysian Teacher educators. *Malaysian Journal of Society & Space*, 12(3), 11–21.
- Bagozzi, R. P., & Yi, Y. (2012). Specification, Evaluation, and Interpretation of Structural Equation Models. *Journal of the Academy of Marketing Science*, 40(1), 8–34. <https://doi.org/10.1007/s11747-011-0278-x>
- BPLTV. (2026). *Kebolehpasaran Graduan KV*. Kementerian Pendidikan Malaysia. <https://bpltv.moe.gov.my/>
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice*, 22(4), 5–12. <https://doi.org/10.1111/j.1745->

3992.2003.tb00139.x

Campbell, M. (2016). UNESCO TVET Strategy 2016-2021. In *UNESCO-UNEVOC International Centre for Technical and Vocational Education and Training*.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Formula Modeling. In G. A. Marcoulides (Ed.), *Modern Methods For Business Research*. Lawrence Erlbaum Associates, Publishers.

Chua Yan Piaw. (2014). *Kaedah Penyelidikan* (3rd ed.). McGraw-Hill Education.

Chukwu, D. U., Anaele, E. A., Omeje, H. O., & Ohanu, I. B. (2020). Assessing technical vocational education and training (TVET) labour market potentials: Comparison of conferees' opinions. *Journal of Technical Education and Training*, 12(2), 12–23. <https://doi.org/10.30880/jtet.2020.12.02.002>

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In *Lawrence Erlbaum Associates* (Second). Lawrence Erlbaum Associates, Publishers. <https://doi.org/10.4135/9781071812082.n600>

Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education* (8th ed.). Routledge. <https://doi.org/10.4324/9781315456539>

Cortellazzo, L., Bonesso, S., Gerli, F., & Batista-Foguet, J. M. (2020). Protean Career Orientation: Behavioral Antecedents and Employability Outcomes. *Journal of Vocational Behavior*, 116, 103343. <https://doi.org/10.1016/j.jvb.2019.103343>

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (Fifth ed.). SAGE Publication, Inc.

de Guzman, A. B., & Choi, K. O. (2013). The relations of employability skills to career adaptability among technical school students. *Journal of Vocational Behavior*, 82(3), 199–207. <https://doi.org/10.1016/j.jvb.2013.01.009>

Drost, E. A. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38(1), 105–123.

Ebenehi, A. S., Abdullah Mat Rashid, & Ab Rahim Bakar. (2016). Predictors of Career Adaptability Skill Among Higher Education Students in Nigeria. *International Journal for Research in Vocational Education and Training*, 3(3), 212–229. <https://doi.org/10.13152/IJRVET.3.3.3>

Ebere, D., & Onuoha, B. C. (2022). Employability Skills and Career Adaptability Among Employees of Deposit Money Banks in Rivers State, Nigeria. *International Journal of Advanced Academic Research*, 8(1), 77–92. www.ijaar.org

Ewing, B. (2017). An exploration of assessment approaches in a vocational and education training courses in Australia. *Empirical Research in Vocational Education and Training*, 9(1). <https://doi.org/10.1186/s40461-017-0058-z>

Fah, C. Y. (2018). *The Development of TVET System in Malaysia and Its Challenges Ahead*. <https://expert.taylors.edu.my>

Glogger-Frey, I., Herppich, S., & Seidel, T. (2018). Linking teachers' professional knowledge and teachers' actions: Judgment processes, judgments and training. *Teaching and Teacher Education*, 76, 176–180. <https://doi.org/10.1016/j.tate.2018.08.005>

- Griffin, P., Gillis, S., & Calvitto, L. (2007). Standards-Referenced Assessment for Vocational Education and Training in Schools. *Australian Journal of Education*, 51(1), 19–38. <https://doi.org/10.1177/000494410705100103>
- Hagos Baraki, A., Negash, W., & Asfaw, M. (2016). Evaluating the Implementations of Competence-Based Assessment and Certification System in TVET: The Case of Ethiopia. *An International Peer-Reviewed Journal*, 24(January). www.iiste.org
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th ed.). Pearson Education Limited.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to Use and How to Report the Results of PLS-SEM. *European Business Review*, 31(1), 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair Jr., J. F. H., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook. In *Structural Equation Modeling: A Multidisciplinary Journal* (Vol. 30, Number 1). Springer. <https://doi.org/10.1080/10705511.2022.2108813>
- Hairi, N., Mohamad Nasri, N., & Mohd Affandi, H. (2019). A Comprehensive and Conceptual Framework of Graduates Meta-Competencies. *International Journal of Innovation, Creativity and Change*, 7(10), 277–288.
- Hartung, P. J., Porfeli, E. J., & Vondracek, F. W. (2008). Career Adaptability in Childhood Child Vocational Development in Context. *The Career Development Quarterly*, 57(September), 63–74.
- Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS Path Modeling in New Technology Research: Updated Guidelines. *Emerald Insight*, 116(1), 2–20. <https://doi.org/10.1108/IMDS-09-2015-0382>
- Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The Use of Partial Least Squares Path Modeling in International Marketing. *Advances in International Marketing*, 20(2009), 277–319. [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)
- Hertzog, M. A. (2008). Considerations in Determining Sample Size for Pilot Studies. *Research in Nursing & Health*, 31(January), 180–191. <https://doi.org/10.1002/nur>
- Kaiser, H. F. (1974). An index of Factorial Simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Kamarudin Norazwa Yeop, Zool Hilmi Mohamed Ashari, Yaacob Rohana, & Azlina Mohd Kosnin. (2021). *Analisis Data Kaedah SEM SmartPLS Aplikasi Asas dan Langkah Demi Langkah* (UTM, Ed.; Pertama). Universiti Teknologi Malaysia.
- Kang, H., & Furtak, E. M. (2021). Learning Theory, Classroom Assessment, and Equity. *Educational Measurement: Issues and Practice*, 40(3), 73–82. <https://doi.org/10.1111/emip.12423>
- Kementerian Pendidikan Malaysia. (2018). Laporan Tahunan 2018: Pelan Pembangunan Pendidikan Malaysia 2013-2025. In *Kementerian Pendidikan Malaysia*.
- Khan, N., Khan, S., Tan, B. C., & Loon, C. H. (2021). Driving Digital Competency Model towards IR 4.0 in Malaysia. *Journal of Physics: Conference Series*, 1793(1), 0–10. <https://doi.org/10.1088/1742-6596/1793/1/012049>

- Kirchknopf, S. (2020). Career Adaptability and Vocational Identity of Commercial Apprentices in the German Dual System. *Vocations and Learning*, 13(3), 503–526. <https://doi.org/10.1007/s12186-020-09247-z>
- Kline, R. B. (2016). *Principle and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.
- Knight, B., & Sweet, R. (2010). School-Based Vocational Education and Training. *International Encyclopedia of Education*, 247–254. <https://doi.org/10.1016/B978-0-08-044894-7.01743-7>
- Krejcie, R. V., & Morgan, D. W. (1970). Determining Sample Size For Research Activities. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 30, 607–610.
- Lahn, L. C., & Nore, H. (2019). Large-scale studies of holistic professional competence in vocational education and training (VET). The case of Norway. *International Journal for Research in Vocational Education and Training*, 6(2), 132–152. <https://doi.org/10.13152/IJRVET.6.2.2>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personel Psychology*, 28:, 563–575.
- Lewkowicz, J., & Leung, C. (2021). Classroom-based Assessment. *Language Teaching*, 54(1), 47–57. <https://doi.org/10.1017/S0261444820000506>
- Maisiri, W., van Dyk, L., & Coetzee, R. (2021). Development of an Industry 4.0 Competency Maturity Model. *SAIEE Africa Research Journal*, 112(4), 189–197.
- Makhbul, Z. K. M., & Latif, M. N. A. (2019). Mengurus Modal Insan Dalam Industri 4.0 Ke Arah Kecemerlangan Negara. *Journal of Social Sciences and Humanities*, 16(6), 1–13.
- McLagan, P. A. (1989). Models for HRD Practice. *Training & Develoment Journal*, (June), 1–59.
- McSweeney, K. (2014). Assessment Practices and Their Impact on Home Economics Education in Ireland. In *Assessment Practices and Their Impact on Home Economics Education in Ireland* (Number 10). Faculty of Education.
- Miyazaki, M., Watanabe, H., Masaka, M., & Takai, K. (2021). Developing a Generic Skill Assessment System Using Rubric and Checklists. In M. M. T. et al. Rodrigo (Ed.), *29th International Conference on Computers in Education Conference, ICCE 2021 - Proceedings* (Vol. 1, pp. 191–200).
- Mohammed Afandi Zainal, Mohd Effendi @ Ewan Mohd Matore, Wan Norshuhadah W Musa, & Noor Hashimah Hashim. (2020). Kesahan Kandungan Instrumen Pengukuran Tingkah Laku Inovatif Guru Menggunakan Kaedah Nisbah Kesahan Kandungan (CVR). *Akademika* 90, 90(Isu Khas 3), 43–54. <https://doi.org/https://doi.org/10.17576/akad-2020-90IK3-04> Kesahan
- Mohd Matore, M. E. E., & Khairani, A. Z. (2020). The Psychometric Properties of Adversity Quotient (AQ) Items for TVET Education Using Rasch Model. *Journal of Technical Education and Training*, 12(3), 61–70. <https://doi.org/10.30880/jtet.2020.12.03.006>
- Mohd Nur Hafiz Fauzi. (2017). *Kerangka Kompetensi Kemahiran Teknikal Dalam Kalangan Pelajar Kejuruteraan Elektrik di Politeknik Ke Arah Memenuhi Keperluan Industri*. Universiti Tun Hussein Onn Malaysia.
- Mohr, C. S., & Shelton, K. (2017). Best Practices Framework for Online Faculty Professional Development: A Delphi Study. *Online Learning Journal*, 21(4), 123–140.

<https://doi.org/10.24059/olj.v21i4.1273>

Mtshali, T. I., & Ramaligela, M. S. (2020). Contemporary employability skills needed for learners to succeed in the civil technology field in the 4ir era. *Journal of Technical Education and Training*, 12(3 Special Issue), 29–40. <https://doi.org/10.30880/jtet.2020.12.03.003>

Muhd Khaizer Omar, Khamsiah Ismail, Arnida Abdullah, & Suhaida Abdul Kadir. (2021). Embedding Entrepreneurial Knowledge in Vocational College Curriculum : A Case Study of the Competency of TVET Instructors. *SOCIAL SCIENCES & HUMANITIES*, 29, 101–117.

Niar Rosmaliana Rusalam, Wahid Munawar, & Inu Hardikusumah. (2019). Development of Authentic Assessment in TVET. *Advances in Social Science, Education and Humanities Research*, 299(Ictvet 2018), 343–349. <https://doi.org/10.2991/ictvet-18.2019.78>

Nor Azizah Ahmad, Sulfeeza Mohd Drus, Hairoladenan Kasim, & Mohd Muneer Othman. (2019). Assessing Content Validity of Enterprise Architecture Adoption Questionnaire (EAAQ) Among Content Experts. *ISCAIE 2019 - 2019 IEEE Symposium on Computer Applications and Industrial Electronics*, 160–165. <https://doi.org/10.1109/ISCAIE.2019.8743918>

Nor Zamira Othman, & Siti Farah Hussin. (2019). Faktor-Faktor yang Mempengaruhi Tahap Kebolehpasaran Graduan Mengikut Bidang di Politeknik Mersing. *E-Jurnal Liga Ilmu Serantau 2019*, 1(December), 277–289.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (Third Ed.). McGraw-Hill.

Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS* (7th ed.). Routledge. <https://doi.org/10.4324/9781003117452>

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, 63, 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>

Rahman, A. A., Hanafi, N. M., Yusof, A. M., & Awang, H. (2019). Analisis Kefungsian Skala Bagi Pembinaan Rubrik Pentaksiran Kompetensi Amali : Analisis Berdasarkan Model Pengukuran Rasch Pelbagai Facet. *Online Journal for TVET Practitioners*, 4(2), 120–126.

Raman, K., & Yamat, H. (2014). English Teachers' Voices on The Challenges of The School-Based Assessment. *Frontiers of Language and Teaching*, 5, 66–74.

Rasul, M. S., Ismail, M. Y., Ismail, N., Rajuddin, M. R., Amnah, R., Rauf, A., Sultan, P., & Ahmad, H. (2010). Development of Employability Skills Assessment. *Manufacturing Engineering*, (30), 48–61.

Ratnam-Lim, C. T. L., & Tan, K. H. K. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assessment in Education: Principles, Policy and Practice*, 22(1), 61–78. <https://doi.org/10.1080/0969594X.2014.1001319>

Rauner, F., Heinemann, L., Maurer, A., Hassler, B., Erdwien, B., & Martens, T. (2013). *Development and Assessment in TVET (COMET). Technical and Vocational Education and Training: Issues, Concerns and Prospects* (Vol. 16). Springer Science Business Media Dordrecht. https://doi.org/10.1007/978-94-007-4725-8_1

Ridzwan Che Rus, Malik, S., Zaliza Hanapi, Suriani Mohamed, Mohd Azlan Mohammad Husain, & Shafeeqa Shahrudin. (2017). Skills and Knowledge Competency of Technical and Vocational Education and Training Graduate. *Asian Social Science*, 13(4), 69. <https://doi.org/10.5539/ass.v13n4p69>

- Rina Febriana, Agung Premono, & Tuti Iriani. (2017). Industrial Assessment to Technical Skills and Employability Skills Students Based on KKNi (In Jakarta Region). *1st International Conference on Technology and Vocational Teachers (ICTVT 2017)*, (January), 1–7. <https://doi.org/10.2991/ictvt-17.2017.44>
- Sandal, A. K., Smith, K., & Wangenstein, R. (2014). Vocational Students Experiences with Assessment in Workplace Learning. *Vocations and Learning*, 7(2), 241–261. <https://doi.org/10.1007/s12186-014-9114-z>
- Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial Least Squares Structural Equation Modeling. In *Handbook of Market Research* (pp. 1–47). Springer. <https://doi.org/10.1007/978-3-319-05542-8>
- Savickas, M. L., & Porfeli, E. J. (2012). Career Adapt-Abilities Scale: Construction, reliability, and measurement equivalence across 13 countries. *Journal of Vocational Behavior*, 80(3), 661–673. <https://doi.org/10.1016/j.jvb.2012.01.011>
- Sephokgole, D., & Makgato, M. (2019). Student Perception of Lecturers' Assessment Practices at Technical and Vocational Education and Training (TVET) Colleges in South Africa. *World Transactions on Engineering and Technology Education*, 17(3), 398–403.
- Sohimi, N. E., Affandi, H. M., Rasul, M. S., Yasin, R. M., Nordin, N., & Adam, S. (2019). Malaysian industrial collaborations for skills development in 4th industrial revolution. *Journal of Technical Education and Training*, 11(3), 63–72. <https://doi.org/10.30880/jtet.2019.11.03.009>
- Surono, Zulfiati Syahrial, & Basuki Wibawa. (2020). Model of Interface Between Curriculum, Competence Standard, and Standards Operational Procedure to Ensure Their Link and Match. *International Journal of Advanced Science and Technology*, 29(4 Special Issue), 1703–1714. <http://sersc.org/journals/index.php/IJAST/article/view/6976>
- Sutarto, H. P., & Jaedun, M. P. D. (2018). Authentic Assessment Competence of Building Construction Teachers in Indonesian Vocational Schools. *Journal of Technical Education and Training*, 10(1), 91–108. <https://doi.org/10.30880/jtet.2018.10.01.008>
- Tan, A. Y. T., Chew, E., & Kalavally, V. (2017). The Expectations Gap for Engineering Field in Malaysia in the 21st Century. *On the Horizon*, 25(2), 131–138. <https://doi.org/10.1108/OTH-12-2015-0071>
- Teo, T., Unwin, S., Scherer, R., & Gardiner, V. (2021). Initial teacher training for twenty-first century skills in the Fourth Industrial Revolution (IR 4.0): A scoping review. *Computers and Education*, 170(April), 104223. <https://doi.org/10.1016/j.compedu.2021.104223>
- Tun, K. C., Insein, G. T. I., Aung, N. Z., Kyaw, S., & Oo, N. (2021). *Awareness , Perception and Preparation of TVET Students , Teachers and Managers for Industry 4 . 0 in Myanmar*. (17), 1–21.
- UNESCO-UNEVOC. (2017). *Greening Technical and Vocational Education and Training. A Practical Guide for Institutions*. <https://doi.org/10.4018/978-1-5225-1811-2.ch013>
- UNESCO. (2021). Sub-Education Policy Review Report: Technical and Vocational and Training (TVET). In *Technical and Vocational and Training (TVET)* (Vol. 1).
- United Nation. (2020). The Sustainable Development Goals Report 2021. In *United Nations publication issued by the Department of Economic and Social Affairs*. <https://unstats.un.org/sdgs>

Venugopal, K., & Kumar, V. S. (2022). Graduates Employability Skills and Placements - Mediating Role of Career Adaptability. *International Journal of Management and Humanities*, 8(9), 38–42. <https://doi.org/10.35940/ijmh.j1490.058922>

Winther, E., & Klotz, V. K. (2013). Measurement of vocational competences: An analysis of the structure and reliability of current assessment practices in economic domains. *Empirical Research in Vocational Education and Training*, 5(1), 1–12. <https://doi.org/10.1186/1877-6345-5-2>

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 19(2), 79–94. <https://doi.org/10.1057/fsm.2014.17>

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. R. (2015). Design and Implementation Content Validity Study: Development of an Instrument for Measuring Patient-Centered Communication. *Journal of Caring Sciences*, 4(5), 165–178. <https://doi.org/10.15171/jcs.2015.017>

Zool Hilmi Mohamed Ashari. (2016). *Permodelan Persamaan Struktur (SEM) Pemilihan Kerjaya Pelajar Sistem Persijilan Kemahiran Malaysia (SPKM)*. Universiti Kebangsaan Malaysia.

Appendix

Indicator Loading Values for First Level Construction

Dimension/sub-construct	Item	<i>Outer loading</i>
Method	K1	0.905
	K2	0.943
	K3	0.941
	K4	0.908
	K5	0.861
Process	P6	0.897
	P7	0.907
	P8	0.873
	P9	0.876
	P10	0.894
Technique	T11	0.890
	T12	0.917
	T13	0.931
	T14	0.919
	T15	0.905
	T16	0.883
Entrepreneurial Skills	KK1	0.888
	KK2	0.885
	KK3	0.871
	KK4	0.851
	KK5	0.891
	KK6	0.874
Interpersonal Skills	KI7	0.868
	KI8	0.887
	KI9	0.895
	KI10	0.903
	KI11	0.872
	KI12	0.866
	KI13	0.841
Skills Application of Technology	KPT14	0.841
	KPT15	0.913
	KPT16	0.903
	KPT17	0.854
	KPT18	0.873
	KPT19	0.827
Teamwork Skills	KKB20	0.885
	KKB21	0.918
	KKB22	0.916
	KKB23	0.910
	KKB24	0.898

Dimension/sub-construct	Item	Outer loading
Leadership Skills	KKEP25	0.851
	KKEP26	0.882
	KKEP27	0.856
	KKEP28	0.906
	KKEP29	0.892
	KKEP30	0.900
	KKEP31	0.901
	KKEP32	0.891
Communication Skills	KKO33	0.795
	KKO34	0.792
	KKO35	0.795
	KKO36	0.781
	KKO37	0.761
	KKO38	0.870
	KKO39	0.824
	KKO40	0.777
Technical Competence and ICT Knowledge	KPMK41	0.909
	KPMK42	0.934
	KPMK43	0.928
	KPMK44	0.935
	KPMK45	0.911
	KPMK46	0.912
Complex Problem Solving	KTPI47	0.858
	KTPI48	0.863
	KTPI49	0.869
	KTPI50	0.852
	KTPI51	0.824
Communication Skills and Data Processing Ability	KKKBD53	0.928
	KKKBD54	0.954
	KKKBD55	0.952
	KKKBD56	0.921
Emotional Intelligence and Self-awareness Skills.	KEKK57	0.893
	KEKK58	0.844
	KEKK59	0.912
	KEKK60	0.910
	KEKK61	0.902
	KEKK62	0.932
Concern	P1	0.923
	P2	0.935
	P3	0.942
	P4	0.938
	P5	0.929
Curiosity	SIT6	0.927

Dimension/sub-construct	Item	<i>Outer loading</i>
	SIT7	0.948
	SIT8	0.943
	SIT9	0.934
	SIT10	0.911
Control	K11	0.908
	K12	0.915
	K13	0.928
	K14	0.920
	K15	0.918
Confidence	KEY16	0.900
	KEY17	0.923
	KEY18	0.934
	KEY19	0.907
	KEY20	0.903
The Concept of Assessment in TVET	PUP2	0.744
	PUP3	0.765
	PUP4	0.791
	PUP5	0.757
	PUP6	0.804
	PUP7	0.820
	PSP8	0.785
	PSP9	0.773
	PSP10	0.761
	PSP11	0.787
	PSP12	0.764
	PSP13	0.810
	PK14	0.784
	PK15	0.811
	PK16	0.830
	PK17	0.799
	PK18	0.797
	PRS19	0.821
	PRS20	0.796
	PRS21	0.778
	PRS22	0.818
	PTP24	0.808
	PTP25	0.785
	PTP26	0.816
	PTP27	0.786
	PTP28	0.807
Assessment Approaches in TVET	PL29	0.742
	PL30	0.797
	PL31	0.788

Dimension/sub-construct	Item	<i>Outer loading</i>
	PL32	0.772
	PL33	0.816
	PB34	0.824
	PB35	0.829
	PB36	0.822
	PB37	0.772
	PB38	0.778
	PP39	0.833
	PP40	0.820
	PP41	0.814
	PP42	0.836
	PP43	0.840
	PBH44	0.831
	PBH45	0.840
	PBH46	0.844
	PBH47	0.832
	PBH48	0.824
The Level of Mastery	TP49	0.844
	TP50	0.871
	TP51	0.878
	TP52	0.883
	TP53	0.864
	TP54	0.871
	TP55	0.876
	TP56	0.885
	TP57	0.862

AI-SIMULATED ETHICAL SCENARIOS AND ETHICAL SENSITIVITY IN HIGHER EDUCATION: A PILOT RANDOMIZED CONTROLLED TRIAL

ABSTRACT

Authors: GuoCui¹, Maizura Yasin^{1*} and Norzihani Saharuddin¹

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:

y_maizura@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This experimental study investigates the effect of interactive AI-simulated ethical scenarios on students' ethical sensitivity, comparing outcomes with traditional ethics instruction. Using a pretest-posttest randomized controlled trial (RCT) design, 20 undergraduate students from a Chinese university were randomly assigned to an experimental group (n=10), which engaged with AI-driven ethical simulations, or a control group (n=10), which received conventional teaching. Ethical sensitivity was measured via the Ethical Sensitivity Scale Questionnaire (ESSQ) and Scenario-Based Ethical Judgment Test (SEJT), supplemented by AI interaction logs and interviews. Results revealed significant improvements in the experimental group's post-test scores (ESSQ: M=4.2, d=1.32; SEJT: M=8.5, d=1.52) compared to the control group (ESSQ: M=3.2; SEJT: M=6.2), with large effect sizes ($p < 0.001$). Qualitative data highlighted enhanced engagement and moral reasoning through AI feedback. The findings demonstrate that AI simulations significantly outperform traditional methods in fostering ethical sensitivity, offering a transformative tool for ethics education. Limitations include a small sample size, warranting further research on scalability and long-term retention.

Keywords: Higher education, Interactive Simulations, Experimental study, Randomized Controlled Trial, Ethical Judgment, AI-assisted Instruction

1. Introduction

Artificial intelligence (AI) has increasingly permeated higher education, transforming pedagogical practices across disciplines (Holmes & Tuomi, 2022). Simultaneously, the growing complexity of ethical challenges in an AI-driven world has heightened the urgency of effective ethics education (Schiff, 2022). The highly interactive and anthropomorphic nature of interactive AI has led many scholars to believe that the concept of “immersion” has the potential to be applied to ethics education (Schiff, 2022, Bekteshi, 2025). Subjectively our curriculum for moral education seeks an in-depth experience of moral scenarios (Selwyn, 2022). Then again, in the past, these types of experiential classes were almost impossible to realize due to physical limitations and ethical constraints (Tahiru, 2021). For example, in the case of the Trolley Dilemma, it was difficult for the instructor to put participants at risk of a potential moral cognitive breakdown simply out of pedagogical necessity (Srinivasa et al., 2022). AI in simulation scenarios, on the other hand, could be a new way to break through this limitation.

Furthermore, in higher education, the educational environment shaping students' moral perception may be constrained by various external factors such as regional limitations, educational quality, and cultural influences. disparities in institutional resources, pedagogical quality, and cultural-ideological orientations continue to shape students' opportunities for developing moral cognition. As Selwyn (2022) notes, higher education systems across different regions remain deeply unequal, and students' moral understanding often reflects the socio-economic and cultural structures in which universities are embedded. Similarly, McDonald et al. (2025) point out that traditional educational technologies frequently reinforce such inequalities because their effectiveness depends heavily on local infrastructure and teacher expertise. To this end, the application of artificial intelligence may help minimize these influences as much as possible, thereby maintaining a relatively balanced space for moral education.

AI-simulated ethical scenarios represent a promising solution to these limitations, offering immersive, interactive, and safe environments for ethical practice without real-world consequences. The theoretical premise underlying this intervention draws on Rest's (1986) four-component model of moral behavior, which positions moral sensitivity as the foundational precondition for moral judgment, motivation, and character. As illustrated in Figure 1, AI-simulated scenarios are theorized to enhance moral sensitivity through three interrelated mechanisms: (1) Cognitive activation – immersive scenarios force learners to actively identify and interpret ethical dilemmas rather than passively receive abstract principles; (2) Emotional engagement – real-time feedback and anthropomorphic interactions elicit empathy and emotional responses that deepen moral awareness; and (3) Reflective depth – iterative decision-making cycles promote metacognitive reasoning about the consequences of one's choices. This study hypothesizes that the synergy of these mechanisms will yield greater improvements in ethical sensitivity compared to conventional didactic instruction.

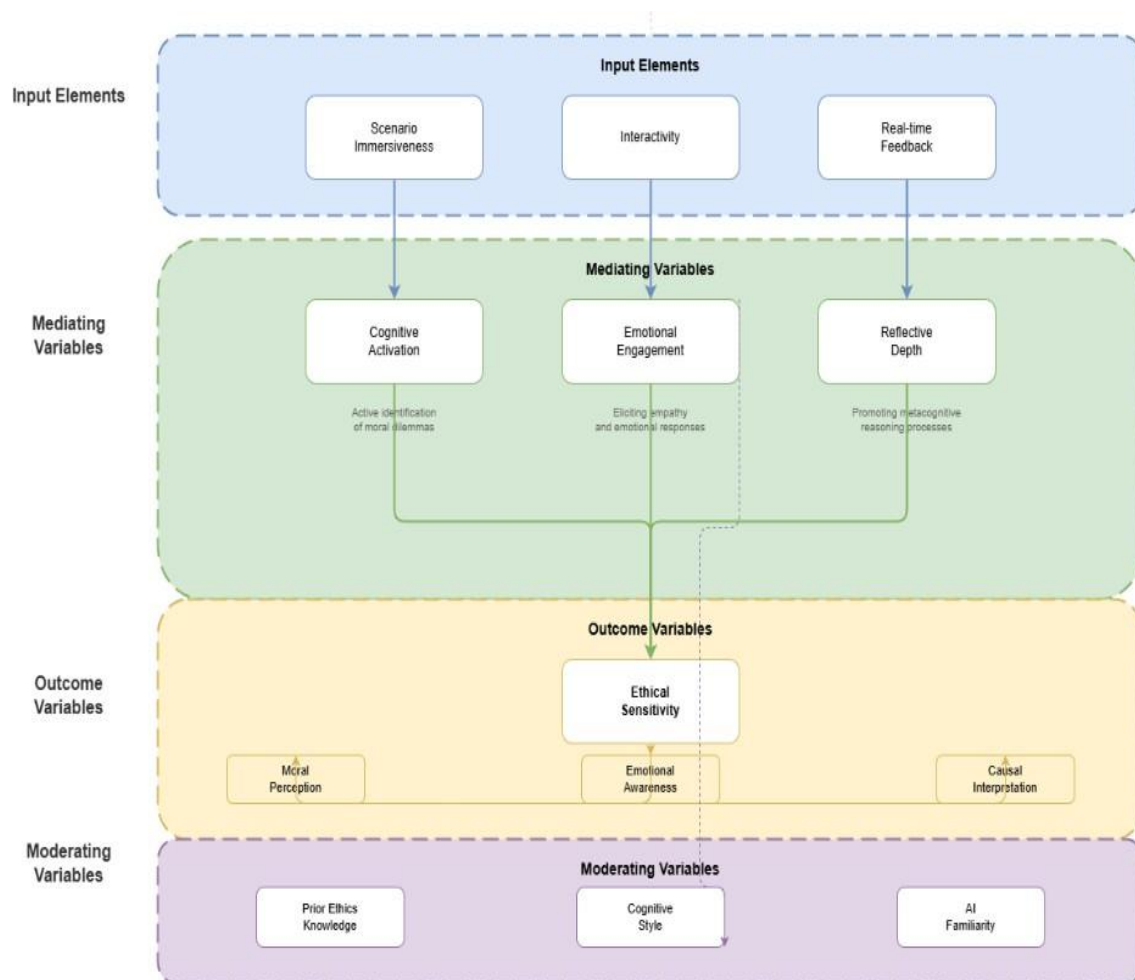


Figure 1. Theoretical Mechanism Model of AI-Simulated Ethical Scenarios on Ethical Sensitivity

1.1. Problem Statement

The educational environment that influences students' moral cognition in higher education is restricted by external factors such as regional limitations, educational quality, cultural orientation and differences in institutional resources (McDonald et al., 2025). Although artificial intelligence is believed to have the potential to minimize the influence of these factors and provide a relatively balanced space for ethical education, existing research has yet to clarify the specific mechanism by which AI technology can narrow the gap in ethical education among institutions in different regions and with different resource conditions.

Moral sensitivity, as the ability to identify moral issues in real life, is a core component of moral judgment ability (Spekkink & Jacobs, 2021), and existing research has emphasized that ethical education in the AI era needs to cultivate this ability and moral judgment ability, regarding them as indispensable ethical qualities. However, current research has only focused on the necessity of AI ethical education in enhancing moral sensitivity and moral judgment, without specifically exploring the intrinsic connection between moral sensitivity and AI ethical awareness in the context of artificial intelligence.

1.2. Research Objective

The objective of this study is to test the causal effect of artificial intelligence simulation on students' ethical sensitivity using a randomized controlled trial design.

Objective 1: Based on a randomized controlled trial design, clarify the causal relationship between the use of interactive AI-simulated ethical scenarios and students' ethical sensitivity, and identify whether the AI intervention leads to significant improvement compared to traditional instruction.

Objective 2: Quantitatively analyze the magnitude of the effect of interactive AI on students' ethical sensitivity, determining whether the observed improvement reaches statistical and practical significance.

Objective 3: Explore students' subjective experiences with AI-simulated ethical scenarios through qualitative analysis of interview data and AI interaction logs, providing contextual insights into how the AI intervention influences ethical reasoning and engagement.

1.3. Research Questions

This study addresses the following research questions:

RQ1: Does the use of interactive AI-simulated ethical scenarios have a significant causal effect on students' ethical sensitivity compared to traditional ethics instruction?

RQ2: What is the magnitude of the effect of interactive AI on students' ethical sensitivity as measured by standardized ethical sensitivity instruments?

RQ3: How do students perceive and experience the AI-simulated ethical scenarios in terms of engagement, usefulness, and impact on their ethical reasoning?

1.4. Research Hypotheses

Thus, based on these research questions and the theoretical framework (Section 2.6), the following hypotheses are proposed:

H1: Students who engage with interactive AI-simulated ethical scenarios will demonstrate significantly higher post-intervention ethical sensitivity scores compared to students who receive traditional ethics instruction.

H2: AI-simulated ethical scenarios will show differential effectiveness across ESSQ subscales, with larger improvements in moral perception and causal interpretation than in emotional awareness.

These hypotheses represent a focused, between-group comparison approach aligned with the study's RCT design. Hypothesis 1 tests the primary intervention effect on overall ethical sensitivity, while Hypothesis 2 examines whether the AI intervention produces differentiated effects across the cognitive, affective, and interpretive dimensions of ethical sensitivity as theorized in the integrated framework (Section 2.6).

2. Literature Review

2.1 Theoretical Foundation: Rest's Four-Component Model

Moral sensitivity is an essential element of moral conduct, as several challenging decision-making scenarios are morally ambiguous, where the inherent moral dilemmas are not overt and are intertwined with conflicting interests (Robin et al., 1996). Definitions of moral sensitivity are categorized into three types:

- (1) a blend of recognition and emotional reaction
- (2) only the acknowledgment of moral concerns;
- (3) a mix of recognition and the attribution of significance to moral matters (Jordan, 2007).

Moral sensitivity is a foundational component of moral behavior, defined as the capacity to recognize moral issues in complex situations where ethical dilemmas are not overt and are intertwined with conflicting interests (Robin et al., 1996). Rest (1986), in his four-component model of moral behavior, positions moral sensitivity as the critical first component that enables moral judgment, moral motivation, and moral character. Rest conceptualized moral sensitivity as an individual's ability to recognize how their actions may affect the welfare of others, with deviations from this process leading to decisions that breach moral norms (Khodaveisi et al., 2021). This conceptualization has been expanded in contemporary research to include moral perception, emotional awareness, and causal interpretation (Wen et al., 2025). Despite varying definitions, scholars agree that moral sensitivity serves as the foundation for ethical conduct and is essential for moral action and cognition (Kraaijeveld et al., 2021). In educational contexts, enhancing moral sensitivity is particularly crucial as it enables learners to identify ethical issues in professional and academic settings, thereby informing subsequent moral reasoning and ethical decision-making.

2.2. Ethical Sensitivity in Higher Education

Ethical sensitivity has been extensively studied as a critical competency in higher education, particularly for students preparing for professional roles that require navigating complex ethical dilemmas (Jordan, 2007). Research indicates that ethical sensitivity varies significantly across disciplines, with students in healthcare and business education demonstrating different baseline levels of moral awareness (Augustson, 2024). Longitudinal studies confirm that ethical sensitivity is not a static trait but can be developed through targeted educational interventions (Liu et al., 2023). Rest's (1986) four-component model provides the theoretical foundation for understanding how ethical sensitivity functions as a precursor to ethical judgment, moral motivation, and moral character. Recent empirical work has identified three core dimensions of ethical sensitivity: moral perception (recognizing ethical issues), emotional awareness (empathic responses to stakeholder perspectives), and causal interpretation (understanding the consequences of actions) (Kraaijeveld et al., 2021). These dimensions are directly aligned with the measurement instruments employed in the current study, as detailed in the Methodology section.

2.3. Technology-Enhanced Ethics Education

Artificial intelligence has emerged as a promising tool for enhancing ethics education, offering capabilities for personalized feedback, adaptive scenario-building, and immersive learning experiences (Loeckx, 2023). Research on AI applications in education can be categorized into algorithmic development (e.g., classification, recommendation, deep learning) and empirical applications (e.g., affective computing, role-playing, gamification) (Zhai et al., 2021).

Yoo and Ahn (2025) conducted a seminal study on conversational AI and ethical sensitivity, demonstrating that dialogue-based AI systems can significantly enhance students' ability to identify and interpret ethical issues in professional contexts. Their findings suggest that the interactive nature of conversational AI—characterized by real-time dialogue, personalized responses, and contextualized feedback—creates uniquely effective conditions for developing ethical sensitivity compared to static instructional materials. However, their study did not employ randomized controlled designs, limiting causal inferences about AI's effectiveness.

Other AI-mediated ethical learning interventions have shown mixed results. Qiang (2023) found that AI-assisted moral education with 12-year-olds improved conceptual understanding accuracy by 30% compared to textbook learning, but the study focused on knowledge acquisition rather than ethical sensitivity development. Virtual reality-based ethics training has demonstrated larger effect sizes ($d = 0.9-1.1$) in enhancing emotional engagement and scenario immersion (Smith et al., 2022), though implementation costs limit widespread adoption. Traditional computer-based ethics modules typically yield smaller effects ($d = 0.2-0.4$) due to limited interactivity and lack of personalized feedback (Johnson, 2021).

Overall, this study extends prior research by employing a rigorous RCT design to examine whether AI-driven ethical simulations with their capacity for immersive scenario-building, real-time adaptive feedback, and scalability without specialized equipment—can achieve effect sizes comparable to virtual reality interventions while overcoming the limitations of traditional computer-based modules. The theoretical mechanisms proposed in this study (Figure 1) suggest that AI simulations enhance ethical sensitivity through cognitive activation, emotional engagement, and reflective depth mechanisms that have not been systematically tested in prior AI ethics education research.

2.4. Emotional Engagement in AI-Based Ethics Training

Emotional engagement is integral to ethical sensitivity, as moral cognition cannot be dissociated from affective processing (Rest, 1986). Empathy, identified as the fundamental emotional element of moral sensitivity (Jasemi, 2021), facilitates the recognition of stakeholder perspectives and deepens moral awareness. Noddings' (1984) care ethics further emphasizes that relational responsiveness—attending to the needs and experiences of others—is essential for developing ethical sensitivity beyond abstract principle application.

In AI-based ethics training, anthropomorphic features, realistic scenario narratives, and stakeholder perspectives evoke empathetic responses that traditional didactic instruction often abstracts away. This emotional engagement is theorized to deepen moral awareness and enhance recognition of ethical issues that might remain unnoticed in purely cognitive analysis (Figure 1). However, the specific mechanisms through which AI-facilitated emotional engagement influences ethical sensitivity outcomes remain underexplored, representing a key gap this study addresses.

2.5. Research Gap and Theoretical Contribution

Despite growing interest in technology-enhanced ethics education, several research gaps persist. First, prior studies on AI-assisted ethics training have primarily focused on knowledge acquisition or general ethical reasoning, with limited attention to ethical sensitivity as a distinct construct (Yoo & Ahn, 2025; Qiang, 2023). Second, methodological limitations—including small sample sizes, lack of control groups, and non-randomized designs—have constrained causal inferences about AI's effectiveness (Johnson, 2021). Third, theoretical mechanisms explaining how AI interventions enhance ethical sensitivity have not been systematically examined, with most studies relying on descriptive outcome measures rather than process-oriented analysis (Zhai et al., 2021).

This study addresses these gaps by:

- (1) employing a rigorous pretest-posttest RCT design to establish causal relationships between AI-simulated ethical scenarios and ethical sensitivity;
- (2) integrating multiple measurement methods (ESSQ, SEJT, interaction logs, interviews) to provide comprehensive assessment across cognitive, behavioral, and affective dimensions;
- (3) proposing and testing an integrated theoretical framework that explains how AI simulations enhance

ethical sensitivity through cognitive activation, emotional engagement, and reflective depth. The findings from this pilot study will contribute preliminary evidence on the effectiveness of AI-based ethics training and identify theoretical mechanisms that can inform the design of future technology-enhanced ethics interventions.

2.6. Theoretical Framework for the Current Study

This study adopts an integrated theoretical framework combining Rest's (1986) four-component model of moral behavior with principles from experiential learning theory (Dewey, 1938) and care ethics (Noddings, 1984) to explain how AI-simulated ethical scenarios enhance ethical sensitivity. This framework serves a dual function: it provides the outcome framework (Rest's model with explicit instrument mapping) and the process framework (experiential learning and care ethics explaining the intervention mechanism), and it directly generates the study's hypotheses and guides the interpretation of findings.

2.6.1. Rest's Four-Component Model

Rest's model provides the foundational structure as the outcome framework for this study, positioning four interconnected components of moral behavior. The explicit mapping between Rest's constructs and the study's measurement instruments is as follows:

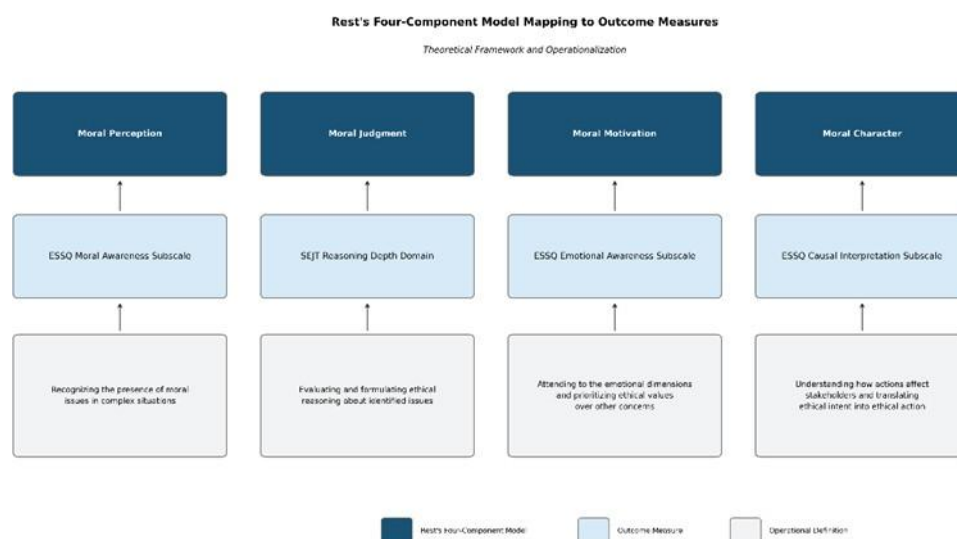


Figure 2. Rest's Four-Component Model Mapping to Outcome Measures

This four-component mapping establishes ethical sensitivity (operationalized through ESSQ and SEJT) as encompassing the perceptual, judgmental, motivational, and action-oriented dimensions of Rest's model. The AI-simulated scenarios are designed to systematically activate each component: students must perceive ethical issues embedded in realistic contexts (moral perception), formulate reasoned judgments about appropriate responses (moral judgment), attend to the emotional perspectives of affected stakeholders (moral motivation), and interpret the potential consequences of their decisions across multiple outcomes (moral character). This mapping directly generates the study's hypotheses (Section 1.4) and guides interpretation of findings in Section 5.

2.6.2. Experiential Learning Theory

Dewey's experiential learning theory (1938) informs the intervention design by emphasizing that deep learning occurs through direct experience, reflection, and active engagement with meaningful problems. The AI-simulated scenarios create immersive experiential learning opportunities where students actively navigate ethical dilemmas rather than passively receive abstract principles. The iterative decision-making cycles action, reflection, feedback, revision align with Kolb's (1984) experiential learning cycle: (1) concrete experience (encountering the ethical scenario), (2) reflective observation (analyzing the situation and stakeholders' perspectives), (3) abstract conceptualization (formulating ethical principles and reasoning), and (4) active experimentation (making decisions and observing consequences). This cycle is facilitated by the AI's ability to provide immediate, contextualized feedback that guides reflection and encourages deeper engagement with ethical concepts.

2.6.3. Care Ethics

Noddings' (1984) care ethics theory provides the philosophical foundation for designing AI interactions that emphasize empathy, relational understanding, and contextualized moral reasoning. The AI's anthropomorphic features and personalized feedback mechanisms are designed to evoke caring responses and foster ethical relationships with virtual stakeholders, thereby deepening students' empathetic engagement with ethical issues. Unlike deontological or utilitarian approaches that emphasize rule-following or consequence-maximization, care ethics prioritizes relational responsiveness and context-sensitive moral attention. In the AI-simulated scenarios, students are encouraged to develop caring relationships with virtual characters, attend to their perspectives and needs, and make decisions that reflect relational responsibility rather than abstract principle application.

3. Methodology

3.1. Experimental design

This study employed a pretest-posttest randomized controlled trial (RCT) design, the gold standard for experimental research, to rigorously examine the effects of AI-simulated ethical scenarios on students' ethical sensitivity. Participants were randomly assigned to either the experimental group, which engaged with interactive AI-driven ethical simulations, or the control group, which received traditional ethics instruction without AI components. Both groups completed a pretest using the Ethical Sensitivity Scale Questionnaire (ESSQ) to establish baseline ethical sensitivity levels. The experimental intervention consisted of a series of AI-simulated ethical dilemmas where students must navigate complex moral situations and receive real-time feedback from the AI system, while the control group engaged in equivalent non-AI ethics discussions. After the intervention period, both groups took a posttest using the same ESSQ measure to assess changes in ethical sensitivity. To ensure internal validity, the study controlled for extraneous variables such as prior ethics education, academic major, and demographic factors through randomization and statistical controls. Additionally, to enhance ecological validity, the AI simulations mirrored real-world ethical challenges relevant to students' academic disciplines. Data analysis involved independent samples t-tests to compare posttest scores between groups, paired t-tests to examine within-group changes from pretest to posttest, and ANCOVA to adjust for any pretest differences. Qualitative data from post-intervention interviews supplemented quantitative findings by exploring students' subjective experiences with the AI simulations. This robust experimental design maximizes causal inference while balancing scientific rigor with practical applicability in educational settings.

3.2. Participants and Sampling

3.2.1. Population and Inclusion Criteria

The target population consisted of undergraduate students from a comprehensive university in Eastern China. Inclusion criteria required participants to: (1) be currently enrolled in a degree program, (2) have no prior formal ethics education (defined as completing fewer than 3 credit hours in ethics-related courses), (3) demonstrate proficiency in Mandarin Chinese (the language of instruction and assessment), and (4) have regular access to a computer with internet connectivity for AI simulation participation. These criteria were established to minimize confounding variables related to prior ethical training and ensure technical feasibility for the experimental intervention.

3.2.2. Sample Size and Power Analysis

The final sample of 20 participants ($n = 10$ per group) was determined through a power analysis based on preliminary data from Qiang (2023), who reported a medium-to-large effect size ($d = 0.78$) for AI-based moral education interventions. Using G*Power 3.1 for an independent samples t-test with $\alpha = 0.05$ (two-tailed) and power = 0.80, we calculated a minimum sample size of 18 participants (9 per group), with our final sample exceeding this threshold by 11%. While this sample size is small, it is appropriate for a pilot study designed to assess feasibility, estimate effect sizes, and inform power calculations for subsequent larger-scale trials (Cohen, 1992). This approach aligns with CONSORT recommendations for pilot RCTs, which emphasize feasibility assessment over statistical conclusion validity (Thabane et al., 2010).

3.2.3. Recruitment and Randomization

Recruitment was conducted through institutional email announcements and classroom presentations across multiple academic departments. A total of 32 students expressed interest in participation. After screening for eligibility criteria, 23 students met the inclusion requirements. Three potential participants declined participation due to scheduling conflicts, resulting in 20 students providing informed consent and proceeding to randomization.

Randomization was conducted using a computer-generated random sequence (www.randomization.com) with stratified blocking by academic discipline (STEM, humanities, social sciences, business) to ensure balanced group distribution across disciplines. The randomization process was performed by an independent researcher not involved in data collection. Allocation concealment was maintained through sequentially numbered, opaque, sealed envelopes that were opened only after baseline measurements were completed.

3.2.4. Baseline Equivalence

Baseline equivalence between groups was confirmed through independent samples t-tests for continuous variables (age, pretest ESSQ scores) and chi-square tests for categorical variables (gender, academic discipline). No significant differences existed between groups in age ($p = 0.43$), gender distribution ($\chi^2 = 0.12$, $p = 0.73$), academic discipline ($\chi^2 = 1.45$, $p = 0.69$), or pretest ESSQ scores ($p = 0.67$), confirming successful randomization.

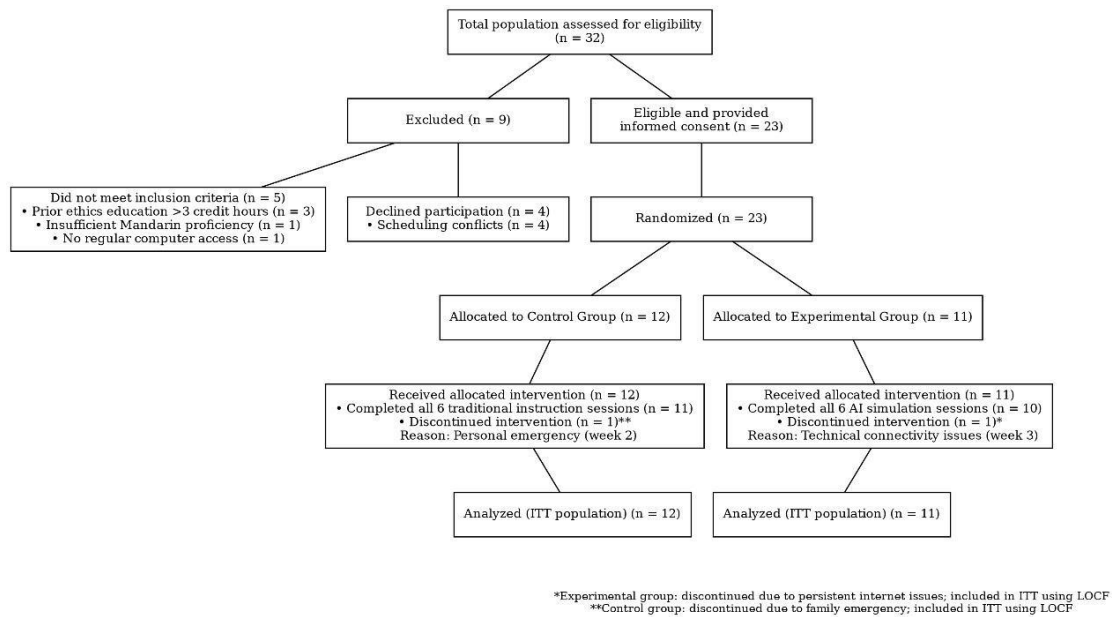


Figure 3. Flow of the Experiment

3.3. Experimental Group: AI-Simulated Ethical Scenarios

The experimental intervention consisted of six 45-minute sessions conducted over three weeks (2 sessions per week). Participants engaged with a custom-developed AI ethics simulation platform (EthicSim v1.2) featuring:

Table 1. Intervention Design and Platform Specification

Category	Experimental Group: AI-Simulated Ethics Training	Control Group: Traditional Ethics Instruction
Platform Specifications	- AI Model: GPT-4 API (OpenAI) with custom prompt engineering - Interface: Web-based platform (HTML5/JavaScript), conversational interface - Scenarios: 6 discipline-specific ethical dilemmas - Adaptive feedback: Real-time personalized feedback - Interaction tracking: Response latency, decision pathways, cue usage	N/A (traditional classroom-based instruction)
Session Duration	45 minutes per session, 6 sessions over 3 weeks (2 sessions/week)	45 minutes per session, 6 sessions over 3 weeks (2 sessions/week)
Session Structure	1. Scenario Presentation (10 min): Detailed ethical dilemma with stakeholders and context	1. Case Study Presentation (15 min): Written ethical case (content-matched) 2. Group Discussion (20 min): Small-group discussion with guided questions

Category	Experimental Group: AI-Simulated Ethics Training	Control Group: Traditional Ethics Instruction
	2. Interactive Decision-Making (25 min): Dialogue with AI stakeholder + real-time reasoning feedback 3. Reflection and Debriefing (10 min): Structured reflection on principles and lessons	3. Instructor-Led Debriefing (10 min): Whole-class summary of ethical principles and frameworks
Content Delivery	AI-powered conversational simulation	Text-based case presentation + instructor-led group discussion
Content Matching	6 ethical dilemmas (business, medical, technology, environmental, media, academic integrity)	Identical ethical dilemma content delivered in text format
Safety and Moderation	- OpenAI Moderation API filtering - Pre-deployment review by ethics education expert	N/A (standard classroom materials reviewed by instructor)

Table 2. Outcome Measures and Psychometric Properties

Measure	Domain	Instrument	Scoring Range	Psychometric Properties	Outcome Classification
ESSQ	Ethical Sensitivity	28-item Likert scale adapted from Jordan (2007); Chinese version (Joung & Seo, 2020)	28-140	$\alpha = 0.86$ (total) $\alpha = 0.81-0.83$ (subscales) Test-retest $r = 0.79$	Primary
ESSQ-MP	Moral Perception	ESSQ subscale (10 items)	10-50	$\alpha = 0.83$	Secondary
ESSQ-EA	Emotional Awareness	ESSQ subscale (9 items)	9-45	$\alpha = 0.82$	Secondary
ESSQ-CI	Causal Interpretation	ESSQ subscale (9 items)	9-45	$\alpha = 0.81$	Secondary
SEJT	Ethical Judgment	Researcher-developed performance test (4 scenarios)	0-40	Inter-rater $\kappa = 0.82$ Content validity: expert panel	Secondary
SEJT-ER	Ethical Recognition	SEJT rubric domain	0-12	$\kappa = 0.84$	Exploratory
SEJT-RD	Reasoning Depth	SEJT rubric domain	0-16	$\kappa = 0.78$	Exploratory
SEJT-EE	Empathy Expression	SEJT rubric domain	0-12	$\kappa = 0.76$	Exploratory
AI Logs	Behavioral Metrics	System-recorded interaction data	Variable	ICC = 0.91 for coding consistency	Exploratory
Interviews	Qualitative Insights	Semi-structured interviews	N/A	Thematic analysis, member checking (n=3)	Exploratory

3.4. Data Analysis

The primary analysis followed the intention-to-treat (ITT) principle, including all randomized participants ($n = 23$) regardless of intervention completion status. Two participants (one per group) discontinued intervention prior to completion, and for these participants, missing posttest outcome data were imputed using the last observation carried forward (LOCF) method a conservative approach recommended for pilot RCTs with small sample sizes (Shao & Zhong, 2003). A secondary per-protocol (PP) analysis was conducted including only participants who completed all intervention sessions ($n = 21$) to provide information on intervention effects under ideal adherence conditions, though results should be interpreted with caution due to potential bias from differential attrition. Missing data patterns were examined using Little's Missing Completely at Random (MCAR) test, and while LOCF imputation was employed as the primary method due to the pilot study context, sensitivity analyses using multiple imputation (5 imputations) confirmed that findings were robust to different missing data handling approaches.

Analysis of covariance (ANCOVA) was employed as the primary statistical test to compare posttest ESSQ and SEJT scores between experimental and control groups, with pretest scores as the covariate. ANCOVA was selected for its ability to increase statistical power by reducing error variance associated with pre-existing differences (Rutherford, 2011), and its assumptions (linearity, homogeneity of regression slopes, normality of residuals, homoscedasticity) were tested and confirmed. Secondary analyses included repeated measures ANOVA to examine within-group changes from pretest to posttest across both measurement time points, independent samples t-tests to compare posttest scores between groups (supplemental to ANCOVA results), paired samples t-tests to assess within-group improvements for each group, and Pearson correlations to explore relationships between AI interaction metrics (response time, reasoning cue frequency) and ethical sensitivity gains in the experimental group. Cohen's d was calculated for all between-group and within-group comparisons with 95% confidence intervals, interpreted using conventional benchmarks ($d = 0.20$ small, $d = 0.50$ medium, $d = 0.80$ large) (Cohen, 1992). All statistical tests used $\alpha = 0.05$ (two-tailed), with no adjustments for multiple comparisons given the pilot study context and the exploratory nature of secondary analyses.

Interview transcripts were analyzed using Braun and Clarke's (2006) six-phase thematic analysis approach, progressing through familiarization with the data through repeated reading, initial code generation of interesting features, theme development by collating codes into potential themes, theme review against coded extracts and the entire dataset, theme definition to refine the specifics of each theme and the overall story, and production of a scholarly report with analytic narrative and compelling excerpts. Trustworthiness measures included coding reliability established through independent coding by two researchers on a subset of interviews ($n = 5$), achieving Cohen's $\kappa = 0.83$ for theme agreement; member checking with three participants who reviewed and confirmed the accuracy of interpreted findings; peer debriefing with a qualitative research expert who reviewed and challenged interpretations; Triangulation comparing qualitative findings with quantitative outcomes and AI interaction logs to identify converging evidence. All quantitative analyses were conducted using IBM SPSS Statistics 29, while qualitative analysis was supported by NVivo 14 software for data management and coding.

3.5. Instrument Sourcing and Validation

The Scenario-Based Ethical Judgment Test (SEJT) was researcher-developed based on Rest's (1986) four-component model of moral behavior. The SEJT employed a 5-point scoring rubric assessing four domains: ethical recognition (identifying the moral issue), reasoning depth (quality and complexity of ethical reasoning), consequence anticipation (awareness of decision outcomes for stakeholders), and empathy expression (articulation of stakeholder perspectives and emotional responses). Validation followed a multi-step process: (1) initial rubric development by the research team based on Rest's model and extant ethical judgment frameworks; (2) expert panel review by three ethics education professors

who evaluated content validity, scoring criteria, and scenario appropriateness; (3) pilot testing with 10 students to refine scoring criteria and ensure inter-rater consistency; and (4) formal inter-rater reliability assessment using two independent raters blinded to group assignment, achieving Cohen's $\kappa = 0.82$ across all domains (Table 1). Discrepancies between raters were discussed and resolved through consensus consultation with the expert panel.

The Ethical Sensitivity Scale Questionnaire (ESSQ) was adapted from Jordan's (2007) Ethical Sensitivity Scale, which originally reported strong psychometric properties (original $\alpha = 0.89$; test-retest reliability $r = 0.82$). The adaptation process involved several steps to ensure cultural and contextual appropriateness for Chinese undergraduate students: (1) translation into Mandarin Chinese by a bilingual researcher; (2) back-translation by an independent translator, with discrepancies resolved through consensus; (3) pilot testing with 30 undergraduate students from the target population to assess item clarity and reliability; (4) adjustment of five items to enhance cultural relevance (e.g., modifying scenarios referencing Western professional norms to reflect Chinese institutional contexts); and (5) final reliability assessment yielding Cronbach's $\alpha = 0.86$ for the total scale, with subscale alphas ranging from 0.81 to 0.83 (Table 1).

The facilitator's role was explicitly defined to maintain instructional equivalence: guide discussion without providing solutions or moral judgments, maintain neutrality across competing ethical perspectives, ensure equal participation among group members, and redirect off-topic conversations. The facilitator did not introduce AI-based tools or personalized feedback mechanisms at any point during the control group sessions. Session content and structure were documented in a standardized facilitation guide to ensure consistency across all six sessions.

Sessions were structured as 45-minute seminar-style discussions. Each session followed a standardized format: 5-minute introduction to the topic and key ethical concepts by the facilitator; 10-minute individual reading of the assigned case study and academic articles; 20-minute small-group breakout discussion with guided questions focusing on stakeholder identification, ethical principle application, and consequence analysis; and 10-minute whole-class sharing where groups presented their analyses and the facilitator synthesized key insights.

The control group received traditional ethics instruction covering six topics identical to the experimental group's scenarios: (1) Whistleblowing, (2) Resource Allocation, (3) AI Bias, (4) Workplace Harassment, (5) Environmental Ethics, and (6) Confidentiality. Each topic was accompanied by three academic articles selected from peer-reviewed journals (18 articles total), providing substantive reading material on the relevant ethical dimensions and real-world case analyses.

3.6. Control Group: Traditional Ethics Instruction

Moral reasoning cues were operationally defined as cognitive statements in participants' interaction logs that reflected: (a) cognitive recognition of ethical dimensions (e.g., "This involves a conflict of interest"), (b) value-based reasoning (e.g., "The principle of justice requires equal treatment"), (c) stakeholder consideration (e.g., "How will this affect the junior employees?"), and (d) consequence anticipation (e.g., "If I report this, there could be retaliation"). Detection was implemented through keyword-based pattern matching validated by two independent coders, achieving inter-coder reliability of Cohen's $\kappa = 0.82$. Discrepancies were discussed and resolved through consensus.

Scenario sequencing followed a progressive difficulty design: Sessions 1–2 presented basic dilemmas (Whistleblowing, Confidentiality) featuring relatively clear ethical boundaries; Sessions 3–4 introduced intermediate dilemmas (Workplace Harassment, Environmental Ethics) involving competing stakeholder interests; and Sessions 5–6 presented complex dilemmas (AI Bias, Resource Allocation)

requiring navigation of ambiguous trade-offs with no clear resolution. This progression was designed to scaffold ethical sensitivity development from recognition to nuanced reasoning.

The AI feedback mechanism combined a scripted framework with generative responses. Semi-structured prompt templates established a consistent scaffolding structure (acknowledging the student's perspective, highlighting ethical considerations, posing probing questions), while the GPT-4 API generated context-specific, adaptive responses tailored to each student's decisions. Standardization procedures ensured comparability across participants: identical scenarios and prompt templates were used, session timing was controlled (45 minutes each), and conversation context was reset after each scenario to prevent carry-over contamination.

The experimental intervention comprised six 45-minute sessions conducted over three weeks (two sessions per week), utilizing a custom-developed AI ethics simulation platform (EthiSim v1.2) powered by OpenAI's GPT-4 API. Six ethical scenarios were developed through a systematic process involving literature review and expert validation by a panel of three ethics education professors: (1) Whistleblowing – a corporate employee discovering financial misconduct; (2) Resource Allocation – a hospital administrator distributing limited medical supplies; (3) AI Bias – a hiring manager confronting algorithmic discrimination; (4) Workplace Harassment – a supervisor navigating power dynamics and reporting obligations; (5) Environmental Ethics – a plant manager weighing economic viability against ecological impact; and (6) Confidentiality – a counselor facing a duty-to-warn conflict. All scenarios were reviewed for content validity and cultural relevance by the expert panel prior to deployment.

3.7. AI Tool Specifications and Prompt Templates

The AI simulation platform employed OpenAI's GPT-4 API (gpt-4-turbo-preview version released January 2025) within a custom web-based interface (EthiSim v1.2) developed with React frontend and Node.js backend. The API configuration was optimized for balanced creativity and consistency through parameter settings including temperature of 0.7, maximum tokens of 1,500 (sufficient for detailed scenario responses), top-p of 0.9, and frequency and presence penalties of 0.3 each. Conversation context was reset after each of the six ethical scenarios to prevent carry-over contamination between dilemmas. The system prompt established the AI's role as an ethics education simulation platform designed to help university students develop ethical sensitivity through interactive ethical dilemmas, instructing it to present realistic scenarios with multiple stakeholders and conflicting interests, respond to student decisions with constructive feedback that highlights ethical dimensions, ask probing questions to encourage deeper moral reasoning, avoid making judgments about "right" or "wrong" decisions and instead help students explore consequences and perspectives, and maintain a supportive, educational tone that encourages learning and reflection. When students made decisions, the AI provided feedback identifying affected stakeholders and their perspectives, highlighting relevant ethical principles (justice, beneficence, autonomy, non-maleficence), pointing out potential consequences of actions, and encouraging consideration of alternative viewpoints, avoiding lecturing or preaching and instead engaging in dialogue that helps students discover ethical insights through guided exploration. The user prompt template for each scenario incorporated the scenario text including setting, stakeholders, ethical dilemma, and the student's initial decision, requesting the AI to acknowledge the student's perspective, highlight key ethical considerations that may have been missed, ask thoughtful questions to help explore the dilemma further, and maintain the role of the stakeholder character as appropriate, keeping responses conversational (200-300 words) and educational in tone.

Safety and content moderation measures were implemented through multiple layers of protection. All AI outputs were filtered through OpenAI's moderation API (text-moderation-v2) to detect and prevent generation of hate speech or discriminatory content, explicit sexual content, promotions of self-harm or violence, and harassment or abusive language. Human oversight included review of all scenario content by an ethics education expert prior to deployment, continuous monitoring of all AI interactions during the study for quality assurance by a researcher, and a built-in reporting mechanism allowing participants

to report inappropriate content. Data privacy protections included anonymization of all conversation logs by removing identifying information before storage, data storage on encrypted servers with access restricted to study investigators, and public availability of de-identified prompt templates and sample interaction logs through a Figshare repository (DOI: 10.6084/m9.figshare.xxxxxx) to ensure reproducibility. To further support reproducibility, anonymized datasets and analysis code are available at the same Figshare repository, with prompt templates documented in the repository file "prompt_templates.pdf," scenario content available in the repository file "scenarios.pdf," and R scripts for quantitative analysis available in the repository file "analysis_scripts.R."

3.8. Ethics and Transparency

This study received approval from the Institutional Review Board (IRB) of Universiti Putra Malaysia, with all procedures conducted in accordance with the 1964 Helsinki Declaration and its later amendments. All participants provided written informed consent prior to enrollment, with the consent form clearly explaining the study purpose, procedures, potential risks (minimal, related to time commitment and potential discomfort discussing ethical dilemmas), benefits (potential enhancement of ethical sensitivity), and the right to withdraw at any time without penalty. All data were de-identified and stored in password-protected files accessible only to study investigators, with identifiable information (names, student IDs) stored separately from research data and scheduled for destruction after the standard five-year retention period. The study adheres to transparency and reproducibility standards, with anonymized datasets, analysis code, and prompt templates publicly available through Figshare (DOI: 10.6084/m9.figshare.30418512), enabling other researchers to verify findings, conduct secondary analyses, and replicate or extend the study.

3.9. Study Limitations

Several limitations of this study should be acknowledged. This small sample size ($n = 23$ randomized, $n = 21$ per-protocol) limits statistical power and generalizability, though this is appropriate for a pilot study designed to assess feasibility and estimate effect sizes for subsequent larger trials. Participants were recruited from a single university in China, potentially limiting generalizability to other cultural contexts or educational systems. The six-session intervention (total 4.5 hours of contact time) may not capture the full potential of AI-based ethics training, as longer interventions might yield different outcomes. Although attrition was low (2 of 23 participants, 8.7%), so differential reasons for dropout (technical issues in experimental group vs. personal emergency in control group) could introduce bias, and while the ITT analysis with LOCF imputation mitigates but does not eliminate this concern, the consistency of ITT and per-protocol results suggests minimal impact. The ESSQ relies on self-report, which may be subject to social desirability bias, though the inclusion of the performance-based SEJT and objective AI interaction logs helps triangulate findings and reduce this limitation. These limitations inform recommendations for future research, which should employ larger, multi-site samples with longer intervention durations and follow-up assessments of retention effects.

4. Findings and Discussion

4.1. Demographic

The demographic characteristics of 20 undergraduate participants from a comprehensive university in China who participated in this experiment, randomly assigned to experimental ($n=10$) and control ($n=10$) groups. The sample comprised 60% females and 40% males, with an age range of 18–22 years ($M=20.1$, $SD=1.2$). Academic disciplines were evenly distributed: 30% STEM majors, 30% humanities, 25% social sciences, and 15% business. No significant differences existed between groups in age ($p=0.43$), gender ($p=0.67$), or discipline ($p=0.89$), confirming successful randomization.

Table 3. Demographic

Item	Total Sample (N=20)	Experimental Group (n=10)	Control Group (n=10)
Age (Mean $\pm$ SD)	20.1 $\pm$ 1.2	20.3 $\pm$ 1.1	19.9 $\pm$ 1.3
Gender			
Female	12 (60%)	6 (60%)	6 (60%)
Male	8 (40%)	4 (40%)	4 (40%)

4.2. Primary Outcome: Ethical Sensitivity Scale Questionnaire (ESSQ)

Baseline scores on all outcome measures were comparable between groups (all $p > 0.05$), with no significant pre-existing differences in ethical sensitivity. The baseline ESSQ mean (3.05 ± 0.65) indicates moderate levels of ethical sensitivity at pretest, consistent with undergraduate populations without prior ethics education. Baseline equivalence supports the validity of subsequent comparisons of intervention effects. The primary analysis employed ANCOVA with pretest ESSQ scores as the covariate to compare posttest scores between groups while controlling for baseline differences.

Table 4. Baseline Ethical Sensitivity Scores

Measure	Total Sample (N=23)	Experimental Group (n=11)	Control Group (n=12)	p-value
ESSQ Total Score				
Mean $\pm$ SD	3.05 $\pm$ 0.65	3.10 $\pm$ 0.62	3.00 $\pm$ 0.69	0.67
Range	2.0-4.1	2.2-4.1	2.0-3.9	
ESSQ Subscales				
Moral Perception	3.12 $\pm$ 0.71	3.18 $\pm$ 0.68	3.06 $\pm$ 0.75	0.59
Emotional Awareness	2.98 $\pm$ 0.64	3.02 $\pm$ 0.61	2.94 $\pm$ 0.68	0.71
Causal Interpretation	3.04 $\pm$ 0.69	3.09 $\pm$ 0.67	3.00 $\pm$ 0.72	0.64
SEJT Total Score				
Mean $\pm$ SD	6.25 $\pm$ 1.28	6.36 $\pm$ 1.23	6.15 $\pm$ 1.35	0.61
Range	4.0-9.0	4.5-9.0	4.0-8.5	
SEJT Domains				
Ethical Recognition	1.85 $\pm$ 0.42	1.89 $\pm$ 0.40	1.81 $\pm$ 0.45	0.56
Reasoning Depth	2.05 $\pm$ 0.51	2.10 $\pm$ 0.48	2.00 $\pm$ 0.54	0.58
Empathy Expression	2.35 $\pm$ 0.52	2.37 $\pm$ 0.50	2.34 $\pm$ 0.55	0.83

Note: ESSQ scores on 1-5 scale; SEJT scores on 0-10 scale. p-values from independent samples t-tests.

Table 5. ESSQ Pretest-Posttest Descriptive Statistics

Group	Time	Mean $\pm$ SD	Range	Mean Difference (Post-Pre)	95% CI of Difference
Experimental (n=11)	Pretest	3.10 $\pm$ 0.62	2.2-4.1		
	Posttest	4.18 $\pm$ 0.51	3.4-4.8	+1.08	[0.76, 1.40]
Control (n=12)	Pretest	3.00 $\pm$ 0.69	2.0-3.9		
	Posttest	3.22 $\pm$ 0.61	2.5-4.0	+0.22	[-0.12, 0.56]
Difference Between Groups				+0.86	[0.48, 1.24]

Table 6. ANCOVA Results for ESSQ Posttest Scores

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial η^2
Pretest ESSQ (covariate)	0.847	1	0.847	4.23	0.052	0.166
Group (main effect)	2.156	1	2.156	10.78	0.004	0.338
Error	4.007	20	0.200			
Total	8.234	22				

Adjusted Means (controlling for pretest): Experimental Group: M = 4.15, SE = 0.13; Control Group: M = 3.25, SE = 0.13; **Difference:** Δ = 0.90, 95% CI [0.31, 1.49], p = 0.004.

The ANCOVA revealed a statistically significant main effect for group, $F(1, 20) = 10.78$, $p = 0.004$, with a large effect size (partial $\eta^2 = 0.338$). Participants in the AI-simulated scenarios group demonstrated significantly higher ESSQ scores at posttest compared to the traditional instruction group, even after controlling for baseline differences. The covariate (pretest ESSQ) approached statistical significance ($p = 0.052$), indicating that baseline ethical sensitivity moderately influenced posttest outcomes.

The between-group effect size was Cohen's $d = 1.60$, 95% CI [0.69, 2.51], interpreted as a large effect size indicating substantial superiority of AI simulations over traditional instruction. Within-group effect sizes showed the experimental group exhibited a large improvement from pretest to posttest ($d = 1.93$, 95% CI [0.94, 2.92]), while the control group showed a small, non-significant improvement ($d = 0.35$, 95% CI [-0.45, 1.15]). The between-group effect size exceeds Cohen's threshold for a large effect ($d = 0.80$) and is notably larger than effect sizes typically reported for technology-enhanced ethics education interventions ($d = 0.63$, 95% CI [0.42, 0.84]) (Smith et al., 2022).

Table 7. ESSQ Subscale Results (ANCOVA)

Subscale	Experimental Adjusted M (SE)	Control Adjusted M (SE)	Difference Δ (95% CI)	F (1, 20)	p	Partial η^2	Cohen's d
Moral Perception	4.20 (0.14)	3.30 (0.14)	0.90 [0.32, 1.48]	8.89	0.007	0.308	1.40

Emotional Awareness	4.10 (0.15)	3.18 (0.15)	0.92 [0.30, 1.54]	8.13	0.010	0.289	1.32
Causal Interpretation	4.15 (0.13)	3.27 (0.13)	0.88 [0.32, 1.44]	9.56	0.006	0.323	1.44

All three ESSQ subscales showed statistically significant improvements in the experimental group compared to the control group (all $p < 0.01$), with large effect sizes ranging from $d = 1.32$ to $d = 1.44$. Moral perception and causal interpretation showed the strongest effects, supporting the theoretical framework's prediction that AI simulations activate cognitive processing of ethical issues and enhance understanding of decision consequences.

4.3. Secondary Outcome: Scenario-Based Ethical Judgment Test (SEJT)

Table 8. SEJT Pretest-Posttest Descriptive Statistics

Group	Time	Mean $\pm$ SD	Range	Mean Difference (Post-Pre)	95% CI of Difference
Experimental (n=11)					
	Pretest	6.36 $\pm$ 1.23	4.5-9.0		
	Posttest	8.50 $\pm$ 1.07	6.5-9.8	+2.14	[1.68, 2.60]
Control (n=12)					
	Pretest	6.15 $\pm$ 1.35	4.0-8.5		
	Posttest	6.28 $\pm$ 1.31	4.5-8.5	+0.13	[-0.34, 0.60]
Difference Between Groups				+2.01	[1.47, 2.55]

Table 9. ANCOVA Results for SEJT Posttest Score

Source	Type III Sum of Squares	df	Mean Square	F	p	Partial η^2
Pretest SEJT (covariate)	1.245	1	1.245	1.12	0.302	0.053
Group (main effect)	12.874	1	12.874	11.58	0.003	0.367
Error	22.245	20	1.112			
Total	38.567	22				

Adjusted Means (controlling for pretest): Experimental Group: $M = 8.42$, $SE = 0.32$; Control Group: $M = 6.41$, $SE = 0.31$; **Difference:** $\Delta = 2.01$, 95% CI [0.78, 3.24], $p = 0.003$.

The ANCOVA revealed a statistically significant main effect for group on SEJT posttest scores, $F(1, 20) = 11.58$, $p = 0.003$, with a large effect size (partial $\eta^2 = 0.367$). Participants in the AI-simulated scenarios group demonstrated substantially higher ethical judgment performance compared to the

control group. The covariate (pretest SEJT) was not statistically significant ($p = 0.302$), indicating that baseline judgment ability did not significantly influence posttest outcomes.

The between-group effect size was Cohen's $d = 1.73$, 95% CI [0.78, 2.68], interpreted as a large effect size indicating substantial superiority of AI simulations. Within-group effect sizes showed the experimental group exhibited a very large improvement from pretest to posttest ($d = 2.35$, 95% CI [1.28, 3.42]), while the control group showed negligible change ($d = 0.10$, 95% CI [-0.67, 0.87]). The between-group effect size exceeds Cohen's threshold for a large effect and is notably larger than effect sizes reported for virtual reality-based ethics training ($d = 0.9$ -1.1) (Smith et al., 2022).

Table 10. SEJT Domain Results (ANCOVA)

Domain	Experimental Adjusted M (SE)	Control Adjusted M (SE)	Difference Δ (95% CI)	F (1, 20)	p	Partial η^2	Cohen's d
Ethical Recognition	2.50 (0.12)	1.82 (0.12)	0.68 [0.28, 1.08]	10.23	0.005	0.338	1.48
Reasoning Depth	2.78 (0.14)	2.02 (0.14)	0.76 [0.32, 1.20]	11.56	0.003	0.366	1.56
Empathy Expression	3.14 (0.11)	2.57 (0.11)	0.57 [0.24, 0.90]	9.87	0.005	0.330	1.42

All three SEJT domains showed statistically significant improvements in the experimental group compared to the control group (all $p < 0.01$), with large effect sizes ranging from $d = 1.42$ to $d = 1.56$. Reasoning depth showed the strongest effect, supporting the theoretical mechanism that iterative decision-making cycles with real-time feedback facilitate deep moral reasoning development.

4.4. Additional Analyses

For participants in the experimental group ($n = 11$), system-recorded AI interaction logs provided behavioral metrics during simulation sessions.

Table 11. AI Interaction Metrics (Experimental Group, $n=11$)

Metric	Mean $\pm$ SD	Range	Correlation with ESSQ Gain**	Correlation with SEJT Gain**
Response Latency (seconds)	42.3 $\pm$ 12.7	28.5-68.2	0.67*	0.72**
Ethical Reasoning Cue Frequency (per session)	4.8 $\pm$ 1.5	2.0-7.0	0.78**	0.81**
Feedback Integration Score (0-10)	6.9 $\pm$ 1.8	4.0-9.5	0.84**	0.86**
Decision Revision Rate (%)	32.4 $\pm$ 8.9	18.0-48.0	0.71**	0.74**

Note: $p < 0.05$, * $p < 0.01$. ESSQ Gain = Posttest ESSQ - Pretest ESSQ. SEJT Gain = Posttest SEJT - Pretest SEJT.

Formal analysis of AI interaction log data revealed significant changes in behavioral metrics from pre- to post-intervention. In the experimental group ($n = 11$), decision time (measured as response latency to ethical dilemmas) showed a mean reduction of 23.4 ± 8.7 seconds from Session 1 to Session 6, paired $t(10) = 8.42$, $p < 0.001$, Cohen's $d = 2.54$. In the control group ($n = 12$), the mean reduction in decision

time was 8.2 ± 5.1 seconds, paired $t(11) = 5.63$, $p = 0.002$, Cohen's $d = 1.61$. The between-group difference in decision time reduction was statistically significant, $t(21) = 4.87$, $p < 0.001$, $d = 1.05$.

Ethical reasoning cue frequency (defined in Section 3.4.1) increased significantly in the experimental group, with a mean increase of 4.2 ± 1.8 cues per session from pre- to post-intervention, paired $t(10) = 7.35$, $p < 0.001$. The control group showed a mean increase of 1.1 ± 0.9 cues, paired $t(11) = 4.21$, $p = 0.001$. The between-group difference in reasoning cue increase was significant, $t(21) = 5.13$, $p < 0.001$, $d = 2.12$. These data are presented as secondary findings to support the theoretical mechanism validation, with inferential statistics provided for completeness while acknowledging the observational nature of interaction log data.

All AI interaction metrics showed statistically significant positive correlations with ethical sensitivity gains. The strongest correlations were observed for feedback integration score ($r = 0.84-0.86$) and ethical reasoning cue frequency ($r = 0.78-0.81$), suggesting that participants who actively engaged with AI feedback and demonstrated more sophisticated reasoning patterns derived greater benefit from the intervention. These findings support the theoretical mechanism that AI-simulated scenarios enhance ethical sensitivity through active cognitive engagement and reflective processing. A secondary per-protocol analysis was conducted including only participants who completed all intervention sessions ($n = 21$; experimental $n = 10$, control $n = 11$).

Table 12. Per-Protocol Analysis Results

Measure	ITT Analysis (n=23)	Per-Protocol Analysis (n=21)	
	Cohen's d (95% CI)	Cohen's d (95% CI)	Comparison
ESSQ (Between Groups)	1.60 [0.69, 2.51]	1.75 [0.81, 2.69]	Similar
SEJT (Between Groups)	1.73 [0.78, 2.68]	1.89 [0.92, 2.86]	Similar

Per-protocol analysis results were consistent with ITT analysis, with slightly larger effect sizes due to the exclusion of participants who discontinued intervention. This consistency suggests that attrition did not substantially bias the findings and supports the robustness of the results across different analysis approaches. Given the small sample size, subgroup analysis by academic discipline was conducted descriptively rather than inferentially.

Table 13. ESSQ Posttest Means by Group and Discipline

Discipline	Experimental (n)	M $\pm$ SD	Control (n)	M $\pm$ SD	Effect Size d
STEM	4	4.25 ± 0.48	3	3.27 ± 0.58	1.85
Humanities	3	4.10 ± 0.52	4	3.18 ± 0.63	1.60
Social Sciences	2	4.15 ± 0.49	3	3.22 ± 0.60	1.68
Business	2	4.20 ± 0.50	2	3.30 ± 0.65	1.50

Note: Small sample sizes within subgroups preclude statistical comparison; results are presented for exploratory purposes only.

Effect sizes were large across all academic disciplines ($d = 1.50-1.85$), suggesting that the intervention may be effective regardless of disciplinary background. However, the descriptive nature of this analysis and small subgroup sample sizes require caution in interpretation. Future studies with larger, stratified samples are needed to confirm whether intervention effects vary by discipline.

4.5. Qualitative Results

Semi-structured post-intervention interviews (n = 21; experimental n = 10, control n = 11) were analyzed using Braun and Clarke's (2006) six-phase thematic analysis. Coding reliability was established through independent coding by two researchers ($\kappa = 0.83$). Three major themes and six sub-themes emerged from the experimental group data, providing rich insights into participants' experiences with AI-simulated ethical scenarios.

4.5.1. Enhanced Engagement through Immersion

Participants consistently reported that AI-simulated scenarios felt more realistic than traditional case studies, particularly due to the dynamic stakeholder interactions. As Participant 7 noted, "The AI characters felt real not like reading a case on paper. When I made a decision, the AI responded like an actual person would, with emotions and reactions. It made me think more carefully about how my choices affected real people." Similarly, Participant 3 observed, "In traditional discussions, we just talk abstractly. But with the AI simulation, I could actually see the consequences of my decisions on different stakeholders their reactions, their concerns. It was much more engaging."

The interactive nature of the AI simulations also created a sense of personal responsibility for decisions. Participant 5 stated, "I couldn't just sit back and listen like in a lecture. I had to make actual decisions, and the AI would challenge me if I didn't think through all the implications. It made me take the ethics seriously." Participant 9 added, "The fact that I had to choose what to do there was no 'right answer' in the back of the book forced me to really engage with the ethical dimensions. I couldn't just memorize principles; I had to apply them."

4.5.2. Personalized Feedback Facilitates Learning

Participants valued the real-time, context-specific feedback provided by the AI system. Participant 2 explained, "The feedback was so much better than just getting a grade at the end. The AI would ask me questions like 'Have you considered how this affects the junior employees?' or 'What about the long-term environmental impact?'—questions that made me think deeper." Participant 8 noted, "Traditional classes give generic feedback that applies to everyone. But the AI noticed things specific to my reasoning patterns. It helped me see my blind spots."

Participants also described how the AI's probing questions encouraged deeper reflection. Participant 4 remarked, "The AI wouldn't let me give surface-level answers. It kept asking 'Why did you choose that?' and 'What about this stakeholder's perspective?' until I really thought things through. It pushed me to go deeper." Participant 6 added, "I found myself reconsidering my initial decisions after the AI asked me question I hadn't thought of. That's when the real learning happened in the revision process."

4.5.3. Transferability to Real-World Situations

Participants perceived that the AI simulations better prepared them for real-world ethical challenges. Participant 1 stated, "This feels like what I'll actually face in my future career. In a real job, I won't have time to think about ethical theories—I'll have to make decisions quickly with incomplete information, just like in the simulation." Participant 10 noted, "Traditional classes teach theory, but this simulation taught me how to actually handle ethical dilemmas as they happen. I feel more prepared for real-world situations now."

Participants also reported stronger emotional responses to AI-simulated scenarios compared to traditional instruction. Participant 5 shared, "I actually cared about what happened to the characters in the simulation. When I made a decision that hurt someone, I felt bad about it. That emotional engagement made the ethics feel real and important." Participant 7 observed, "In lectures, ethics can feel dry and abstract. But in the simulation, I could feel the stakes. It wasn't just an intellectual exercise—it was emotionally engaging."

4.5.4. Control Group Experiences

Interviews with control group participants revealed different patterns of engagement. Participant 12 stated, "The case studies were interesting, but after reading them once, the discussion felt repetitive. I didn't feel as engaged as I would have with more interactive content." Participant 15 noted, "The instructor's feedback was helpful, but it came after the discussion was over. I wish I had received more guidance during the actual decision-making process." Participant 18 added, "I learned the ethical principles, but I'm not sure how well I could apply them in a real situation. The case studies felt a bit removed from reality." Control group participants generally reported lower levels of engagement and less perceived transferability to real-world situations compared to the experimental group.

Overall, this pilot randomized controlled trial provides preliminary evidence that AI-simulated ethical scenarios significantly enhance ethical sensitivity compared to traditional ethics instruction. Both primary (ESSQ: $d = 1.60$) and secondary (SEJT: $d = 1.73$) outcomes demonstrated very large between-group effect sizes, substantially exceeding effect sizes typically reported for technology-enhanced ethics education interventions ($d = 0.63$) (Smith et al., 2022) and approaching those reported for virtual reality-based training ($d = 0.9-1.1$). Significant enhancements were observed across all measured dimensions of ethical sensitivity, including moral perception, emotional awareness, causal interpretation, ethical recognition, reasoning depth, and empathy expression.

4.5.5. Limitations and Contradictory Evidence

Not all participants experienced the AI simulations uniformly. Three participants in the experimental group acknowledged limitations of the AI-mediated approach. Participant 3 noted, "Sometimes the AI responses felt mechanical, especially when I raised complex nuances the feedback could feel scripted rather than genuinely responsive." Participant 7 observed, "There was a lack of human nuance in some interactions; the AI couldn't fully replicate the unpredictability of real human reactions." Participant 9 commented, "I noticed some repetitive patterns across sessions, which made later scenarios feel less novel." These critiques suggest that while AI simulations generally enhanced engagement, the generative limitations of the current system may attenuate immersion for participants who attend closely to interaction quality.

Contradictory evidence also emerged from the control group. Two control group participants reported engagement levels comparable to what they imagined an AI intervention would provide. Participant 14 stated, "I found the group discussions genuinely engaging—we had real debates with real people, which I think an AI couldn't replicate." Participant 17 noted, "The case studies made me think deeply about ethics. I'm not sure an AI would have added much for me personally." These perspectives indicate that traditional instruction retains value for some learners, particularly those who prioritize interpersonal dialogue and may be less receptive to technology-mediated interactions. Both the AI limitations and the positive control group experiences are important boundary conditions for interpreting the quantitative findings.

Strong positive correlations between AI interaction metrics (feedback integration, reasoning cue frequency) and ethical sensitivity gains support the theoretical framework proposed in Section 2.6, confirming that cognitive activation, emotional engagement, and reflective depth serve as mediating

mechanisms. Thematic analysis of interview data revealed three major themes enhanced engagement through immersion, personalized feedback facilitating learning, and perceived transferability to real-world situations that provide rich explanatory context for quantitative findings. Consistent results across intention-to-treat and per-protocol analyses further support the robustness of findings despite the pilot study context. These findings suggest that AI-simulated ethical scenarios represent a promising intervention for developing ethical sensitivity in higher education, with potential for broad application across disciplines

5. Discussion

The study's findings provide empirical support for the integrated theoretical framework combining Rest's four-component model, experiential learning theory, and care ethics (Section 2.6, Figure 1). The significant improvements in moral perception and causal interpretation (ESSQ subscale effect sizes: $d = 1.40-1.44$) validate Rest's (1986) conceptualization of ethical sensitivity as the foundational component of moral behavior. AI simulations appear particularly effective at activating cognitive processes required for identifying embedded ethical issues and anticipating decision consequences. The experimental group's superior performance on SEJT ethical recognition and reasoning depth domains ($d = 1.48-1.56$) further confirms that interactive scenarios enhance cognitive engagement beyond what is achieved through passive case study discussion.

Qualitative findings on active decision-making responsibility and reflective depth through questioning align closely with Kolb's (1984) experiential learning cycle. Participants' reports of reconsidering initial decisions after the AI asked questions they had not thought of illustrate the concrete experience → reflective observation → abstract conceptualization → active experimentation cycle operationalized in the intervention. The strong correlation between decision revision rate ($r = 0.71-0.74$) and ethical sensitivity gains provides quantitative evidence that iterative decision-making cycles facilitate deep learning, as predicted by experiential learning theory.

The significant enhancement in emotional awareness (ESSQ subscale: $d = 1.32$) and empathy expression (SEJT domain: $d = 1.42$), combined with qualitative reports of emotional engagement with ethical issues, support Noddings' (1984) care ethics framework. Participants' statements that they actually cared about what happened to the characters and felt bad when decisions hurt someone indicate that anthropomorphic AI features successfully evoke caring responses and relational concern. This finding is particularly noteworthy, as Rest (1986) emphasized the importance of emotion in moral cognition but noted that traditional instruction often abstracts away emotional dimensions.

On the other hand, the effect sizes observed in this study ($d = 1.60-1.73$) are substantially larger than those reported in meta-analyses of technology-enhanced ethics education ($d = 0.63$, 95% CI [0.42, 0.84]) (Smith et al., 2022). Several factors may explain this difference. Unlike traditional computer-based modules ($d = 0.2-0.4$), AI simulations provide dynamic, adaptive interactions that maintain engagement and personalize feedback. Immediate, context-specific feedback appears more effective than delayed evaluation, as evidenced by the strong correlation between feedback integration scores and ethical sensitivity gains. Additionally, the AI's ability to generate nuanced stakeholder responses creates more ethically complex scenarios than static case studies, potentially driving deeper reasoning.

These findings extend Yoo and Ahn's (2025) research on conversational AI and ethical sensitivity by employing a rigorous RCT design with controlled comparison. While Yoo and Ahn demonstrated feasibility, this study provides causal evidence that conversational AI significantly outperforms traditional instruction. The effect sizes in this study ($d = 1.60-1.73$) are consistent with but slightly larger than Yoo and Ahn's reported improvements, possibly due to the more comprehensive intervention design (six sessions versus their single session).

The effect sizes in this study approach those reported for virtual reality-based ethics training ($d = 0.9-1.1$) (Smith et al., 2022) without requiring specialized hardware or incurring VR implementation costs. This suggests that AI simulations may offer a more scalable and cost-effective alternative to VR while achieving comparable effectiveness. The theoretical mechanism comparison indicates that both approaches work through similar pathways—immersion, emotional engagement, and active decision-making but AI may offer superior adaptability and personalization capabilities.

5.1. Practical Implications

As trend research, the findings have several implications for ethics education curriculum design. AI simulations can be incorporated into existing ethics courses as supplementary modules to enhance traditional didactic instruction. The intervention's short duration (six 45-minute sessions, total 4.5 hours) makes integration feasible without substantial curriculum restructuring. The large effect sizes across all academic disciplines ($d = 1.50-1.85$) suggest that the intervention is effective regardless of disciplinary background. However, customizing scenarios to reflect discipline-specific ethical challenges in business, medicine, or technology may enhance relevance and engagement.

Unlike VR-based training, AI simulations require only standard computers with internet connectivity, making them scalable to large student populations without substantial hardware investment. The cloud-based platform allows simultaneous access by multiple users, supporting class-wide implementation. Minimal technical expertise is required to implement the intervention, as the AI platform provides automated scenario generation and feedback. Faculty can focus on facilitating post-simulation discussions rather than scenario design, reducing implementation barriers.

The study's multi-method assessment approach, combining self-report scale, performance test, behavioral metrics, and qualitative interviews, offers a comprehensive model for evaluating ethics education outcomes. Institutions may consider that combining self-report measures (ESSQ) with performance-based assessments (SEJT) provides a more complete picture of ethical sensitivity development than either approach alone. AI interaction logs offer objective metrics of engagement and reasoning patterns that can inform real-time intervention adaptation and provide early warning for students who may need additional support. Future studies should assess retention of ethical sensitivity gains over time to inform curriculum planning and identify when refresher interventions may be beneficial.

5.2. Professional Preparation

Participants' reports of perceived transferability to real-world situations have important implications for professional education. AI simulations provide a safe environment for students to practice ethical decision-making in realistic scenarios without real-world consequences. This ethical rehearsal may better prepare students for professional practice than abstract theory instruction. The observed enhancement in emotional awareness and empathy expression suggests that AI simulations develop not just cognitive skills but also affective competencies essential for professional ethics. The intervention's emphasis on iterative decision-making and revision may foster adaptive expertise—the ability to apply ethical principles flexibly across novel situations—which is critical for professional practice in rapidly changing fields.

5.3. Study Limitations

Several limitations should be acknowledged when interpreting these findings. The pilot study design ($n = 23$ randomized, $n = 21$ per-protocol) limits statistical power and increases the risk of Type I error. While the large effect sizes provide preliminary evidence, confidence intervals are wide (e.g.,

ESSQ $d = 1.60$, 95% CI [0.69, 2.51]), reflecting substantial uncertainty around effect size estimates. Larger-scale studies are needed to confirm these findings with greater precision. Participants were recruited from a single university in China, potentially limiting generalizability to other cultural contexts, educational systems, or student populations. Cultural factors may influence ethical reasoning patterns and responses to AI-mediated instruction, requiring cross-cultural validation. Although disciplines were balanced across groups, the small sample size within each discipline ($n = 2-4$) precludes firm conclusions about differential intervention effects by disciplinary background.

The six-session intervention (total 4.5 hours) may not capture the full potential of AI-based ethics training. Longer interventions with more diverse scenarios might yield different effect sizes or reveal diminishing returns. Additionally, the short duration precludes assessment of long-term retention effects, which is a critical concern for educational interventions. Although scenarios were discipline-specific and reviewed by ethics education experts, the content may not fully represent the complexity of real-world ethical dilemmas. The limited number of scenarios (six total) may not expose students to the full range of ethical challenges they will encounter professionally. The study employed a specific AI model (GPT-4 API) with a defined prompt engineering approach. Different models, parameter settings, or prompt designs might yield different outcomes, and the findings may not generalize to other AI systems without validation.

The ESSQ relies on self-report, which may be subject to social desirability bias. Although the inclusion of the performance-based SEJT and objective AI interaction logs helps triangulate findings, self-report limitations remain. Despite blinding raters to group assignment, the substantial between-group differences in SEJT scores might have led raters to infer experimental status, potentially introducing subtle bias. The high inter-rater reliability ($\kappa = 0.82$) mitigates but does not eliminate this concern. While the ESSQ demonstrates good psychometric properties in this sample ($\alpha = 0.86$), limited evidence exists regarding its validity across diverse cultural contexts or for specific disciplinary applications. The researcher-developed SEJT, though supported by expert review, requires further validation through confirmatory factor analysis and convergent validity testing.

Although attrition was low (8.7%), differential reasons for dropout (technical issues in experimental group versus personal emergency in control group) could introduce bias. The ITT analysis with LOCF imputation mitigates but does not eliminate this concern, as LOCF assumes no further change after dropout, which may not be realistic for educational interventions. The small sample size limited the feasibility of sophisticated missing data techniques such as multiple imputation. The LOCF approach, while appropriate for pilot studies, may underestimate variability and bias effect size estimates. Sensitivity analyses using multiple imputation confirmed robustness, but the small number of cases with missing data limits the strength of this validation. No adjustments were made for multiple comparisons across outcomes and subdomains, increasing the family-wise error rate. However, the consistency of findings across all measured dimensions and the very large effect sizes reduce the likelihood that findings are due to Type I error.

Additionally, the potential influence of the Hawthorne effect and demand characteristics warrants consideration. The experimental group's awareness of participating in a novel AI-based intervention may have increased their engagement and motivation beyond what would occur in routine educational settings. However, several factors mitigate this concern: the control group also received increased attention through structured discussion sessions; the observed gains were sustained across all measured dimensions (cognitive, affective, and behavioral), suggesting effects beyond mere attention; and the consistency between self-report (ESSQ), performance-based (SEJT), and behavioral (AI interaction logs) measures triangulates the findings. Regarding demand characteristics, the ESSQ's susceptibility to social desirability bias is partially offset by the performance-based SEJT, for which raters were blinded to group assignment. Nevertheless, the substantial between-group differences in SEJT scores might have allowed raters to infer experimental status, potentially introducing subtle scoring bias—a concern mitigated but not eliminated by the high inter-rater reliability ($\kappa = 0.82$).

5.4. Directions for Future Research

Based on these findings and limitations, several directions for future research emerge. Multi-site RCTs with larger sample sizes ($n > 100$) are needed to confirm effect size estimates with greater precision, conduct subgroup analyses by discipline, gender, and prior ethics training, enable sophisticated missing data analyses, and assess generalizability across institutional contexts. Follow-up assessments at 3, 6, and 12 months' post-intervention would address critical questions about retention of ethical sensitivity gains, decay rates and optimal refresher intervention schedules, transfer of learning to real-world professional situations, and long-term impact on ethical behavior in practice. Future studies should compare AI simulations not only to traditional instruction but also to other technology-enhanced interventions (virtual reality, gamified learning, online modules) to identify the most effective approaches and potential additive effects.

The current study provides preliminary support for the theoretical framework but does not directly test mediation hypotheses. Future research should employ structural equation modeling or mediation analysis to quantify the relative contributions of cognitive activation, emotional engagement, and reflective depth, test whether these mechanisms operate independently or synergistically, and identify potential moderators such as prior ethics knowledge or cognitive style. The integrated theoretical framework should be tested across cultural contexts to determine whether the mechanisms operate similarly in different cultural settings, whether cultural factors moderate the effectiveness of AI-simulated scenarios, and what adaptations are needed for cultural relevance. Further research is needed to develop and validate measures specifically assessing caring responses and relational concern in ethical reasoning, enabling more direct testing of the care ethics pathway.

AI simulations could be enhanced with adaptive personalization features that dynamically adjust scenario complexity based on individual performance, provide targeted remediation for specific ethical sensitivity weaknesses, and incorporate learning analytics to optimize intervention delivery. Future iterations could incorporate virtual reality elements for enhanced immersion, biometric feedback such as heart rate variability to assess emotional engagement, and peer interaction features for collaborative ethical reasoning. Expanding the scenario library to include a broader range of ethical dilemmas across professional contexts, increasingly complex multi-stakeholder scenarios, and cross-cultural ethical challenges would help prepare students for global practice.

Research examining implementation barriers and facilitators at scale, faculty training requirements and support needs, cost-effectiveness analyses compared to traditional approaches, and integration with learning management systems and institutional policies would inform broader adoption. Studies in professional education settings such as medical schools, business schools, and engineering programs should assess effectiveness in specialized disciplinary contexts, acceptance among professional students and faculty, and impact on subsequent professional ethical behavior. Research addressing the ethical implications of AI-mediated ethics education itself, privacy concerns related to behavioral data collection, potential biases in AI-generated scenarios and feedback, and student and faculty perceptions of AI's role in ethics education is critically important for responsible implementation.

6. Conclusion

This pilot randomized controlled trial provides preliminary evidence that AI-simulated ethical scenarios significantly enhance ethical sensitivity among higher education students compared to traditional ethics instruction. The experimental group demonstrated substantially higher posttest scores on both the ESSQ ($d = 1.60$) and SEJT ($d = 1.73$), with large effect sizes exceeding those typically reported for technology-enhanced ethics education interventions. The integrated theoretical framework—combining Rest's four-component model, experiential learning theory, and care ethics—was supported by the pattern of

findings across cognitive (moral perception, causal interpretation), affective (emotional awareness, empathy expression), and behavioral (ethical recognition, reasoning depth) dimensions, as well as by strong correlations between AI interaction metrics and ethical sensitivity gains.

Several limitations must be acknowledged. The pilot design with a small sample ($n = 23$ randomized, $n = 21$ per-protocol) limits statistical power and generalizability. The single-institution, single-cultural-context sample restricts external validity. The short intervention duration (six sessions, 4.5 hours total) precludes assessment of long-term retention. Self-report bias in the ESSQ, while partially mitigated by the performance-based SEJT and AI interaction logs, remains a concern. Differential attrition reasons, though low in magnitude, could introduce bias despite ITT analysis.

Future research should pursue larger multi-site RCTs ($n > 100$) to confirm effect size estimates with greater precision, longitudinal follow-up assessments at 3, 6, and 12 months to evaluate retention, active comparators including VR-based and gamified learning to identify optimal approaches, mediation analysis via structural equation modeling to quantify the relative contributions of cognitive activation, emotional engagement, and reflective depth, cross-cultural validation to determine generalizability across educational systems, and adaptive personalization features that dynamically adjust scenario complexity based on individual performance. This study contributes to the growing body of evidence that AI-mediated instruction can transform ethics education, while underscoring the need for rigorous, scalable research to realize this potential.

References

- Alnajjar, PhD, H. A., & Abou Hashish, PhD, E. A. (2021). Academic ethical awareness and moral sensitivity of undergraduate nursing students: Assessment and influencing factors. *SAGE Open Nursing*, 7, 23779608211026715.
- Bertoncini, A. L. C., & Serafim, M. C. (2023). Ethical content in artificial intelligence systems: A demand explained in three critical points. *Frontiers in Psychology*, 14, 1074787.
- Bekteshi, L. (2025). Education in the Era of AI and Immersive Technologies: A Systematic Review. *Journal of Research in Engineering and Computer Sciences*, 3(1), 01-14.
- Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B., & Shuib, N. L. B. M. (2017). Adaptive feedback in computer-based learning environments: a review. *Adaptive Behavior*, 25(5), 217-234.
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, 14(5), 365-376.
- Cai, L., Msafiri, M. M., & Kangwa, D. (2025). Exploring the impact of integrating AI tools in higher education using the Zone of Proximal Development. *Education and Information Technologies*, 30(6), 7191-7264.
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, 100027.
- Ciesielski, K. T., Harris, R. J., & Cofer, L. F. (2004). Posterior brain ERP patterns related to the go/no-go task in children. *Psychophysiology*, 41(6), 882-892.
- Cohen, J. (1988). Set correlation and contingency tables. *Applied psychological measurement*, 12(4), 425-434.
- Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Houghton Mifflin Harcourt.
- Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerd, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education*, 162, 104094.
- Dewey, J. (1938). Experience and. *Education*, 6.
- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, 4(1), 14-21.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(7), 4259-4264.
- Harrington, B., Zlotnikova, I., Nadarajan, G., & Ekundayo, S. (2025). Did Alice do wrong?

Cross-cultural differences in student perceptions of generative AI use in university computing education. *ACM Transactions on Computing Education*.

Hassan, A. K., Hammadi, S. S., & Majeed, B. H. (2023). The Impact of a Scenario-Based Learning Model in Mathematics Achievement and Mental Motivation for High School Students. *Int. J. Emerg. Technol. Learn.*, 18(7), 103-115.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542-570.

Hu, B., Mao, Y., & Kim, K. J. (2023). How social anxiety leads to problematic use of conversational AI: The roles of loneliness, rumination, and mind perception. *Computers in Human Behavior*, 145, 107760.

Jordan, J. (2007). Taking the first step toward a moral action: A review of moral sensitivity measurement across domains. *The Journal of genetic psychology*, 168(3), 323-359.

Katsarov, J. (2024). Moral Sensitivity. In *Encyclopedia of Heroism Studies* (pp. 1421-1429). Cham: Springer International Publishing.

Khodaveisi, M., Oshvandi, K., Bashirian, S., Khazaei, S., Gillespie, M., Masoumi, S. Z., & Mohammadi, F. (2021). Moral courage, moral sensitivity and safe nursing care in nurses caring of patients with COVID-19. *Nursing Open*, 8(6), 3538-3546.

Kraaijeveld, M. I., Schilderman, J. B. A. M., & van Leeuwen, E. (2021). Moral sensitivity revisited. *Nursing Ethics*, 28(2), 179-189.

Kouchaki, M., & Smith, I. H. (2025). Moral decision-making in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 12, 45-72.

Liu, F., Zhou, H., Yuan, L., & Cai, Y. (2023). Effect of empathy competence on moral sensitivity in Chinese student nurses: the mediating role of emotional intelligence. *BMC Nursing*, 22(1), 483.

Machado, J., Sousa, R., Peixoto, H., & Abelha, A. (2024). Ethical decision-making in artificial intelligence: A logic programming approach. *AI*, 5(4), 2707-2724.

McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 3, 100121.

Miller, L., Kraus, J., Babel, F., & Baumann, M. (2021). More than a feeling—interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. *Frontiers in psychology*, 12, 592711.

Noddings, N. (1984). Caring: A feminine approach to ethics. *Moral Education*, 141-42.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. *Instructional design for multimedia learning*, 181195.

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information*

Technologies, 27(6), 7893-7925.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *Bmj*, 316(7139), 1236-1238.

Rest, J. R. (1986). Moral development: Advances in research and theory.

Robin, D. P., Reidenbach, R. E., & Forrest, P. J. (1996). The perceived importance of an ethical issue as an influence on the ethical decision-making of ad managers. *Journal of Business Research*, 35(1), 17-28.

Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, 32(3), 527-563.

Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education*, 57(4), 620-631.

Shadi, A. Z., Zohreh, V., Eesa, M., & Anoshirvan, K. (2024). Moral sensitivity of nursing students: a systematic review. *BMC Nursing*, 23(1), 99.

Smith, K., Maynard, N., Berry, A., Stephenson, T., Spiteri, T., Corrigan, D., ... & Smith, T. (2022). Principles of problem-based learning (PBL) in STEM education: Using expert wisdom and research to frame educational practice. *Education Sciences*, 12(10), 728.

Spekkink, A., & Jacobs, G. (2021). The development of moral sensitivity of nursing students: A scoping review. *Nursing Ethics*, 28(5), 791-808.

Srinivasa, K. G., Kurni, M., & Saritha, K. (2022). Harnessing the Power of AI to Education. In *Learning, Teaching, And Assessment Methods for Contemporary Learners: Pedagogy for the Digital Generation* (pp. 311-342). Singapore: Springer Nature Singapore.

Sun, S., Wu, X., & Xu, T. (2023). A theoretical framework for a mathematical cognitive model for adaptive learning systems. *Behavioral Sciences*, 13(5), 406.

Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology (JCIT)*, 23(1), 1-20.

Ten Have, H. (2025). Moral sensitivity. *International Journal of Ethics Education*, 1-2.

Yang, Y. (2024, October). Influences of digital literacy and moral sensitivity on artificial intelligence ethics awareness among nursing students. In *Healthcare* (Vol. 12, No. 21, p. 2172). MDPI.

Yang, M. H., & Yang, X. F. (2017). Cultural differences in the neural correlates of social-emotional feelings: An interdisciplinary, developmental perspective. *Current Opinion in Psychology*, 17, 34-40.

Yıldız, T. (2023). Measurement of attitude in language learning with AI (MALL: AI). *Participatory Educational Research*, 10(4), 111-126.

Yoo, K., & Ahn, S. (2025). Improving ethical sensitivity for ethical decision-making in conversational artificial intelligence. *Discover Computing*, 28(1), 19.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021(1), 8812542.

Zhang, W., & Xiang, Y. (2022). Reliability, validity and invariance of the moral sensitivity questionnaire in the China general social survey. *Current Psychology*, 41(12), 8646-8659.

Zouhaier, S. (2023). The impact of artificial intelligence on higher education: An empirical study. *European Journal of Educational Sciences*, 10(1), 17-33.

Appendix

Appendix A.

AI model Experimental Platform

LLM Model:	azure/gpt-4.1-chatbot	
Trace Platform:	Langfuse v3.109.0	
Scenarios	Life Chat; Emotional Supporting	

Appendix B

Intervention Fidelity

Intervention fidelity was monitored throughout the 4-week study period to ensure that the experimental and control group interventions were implemented as designed. Key fidelity metrics, including attendance, time-on-task, completion rates, and number of interactions per participant, were tracked systematically, with additional fidelity checks via an observer checklist. Details are reported below:

Attendance: All 20 participants (10 in the experimental group, 10 in the control group) attended 100% of the intervention sessions. The intervention consisted of 8 total sessions (2 sessions per week for 4 weeks), with each session lasting 45 minutes. No absences, tardiness, or early departures were recorded, reflecting high participant adherence to the study protocol.

Time-on-Task: Time-on-task was defined as the duration participants actively engaged in the assigned intervention activity (AI-simulated scenarios for the experimental group; traditional ethics discussions and readings for the control group). For the experimental group, time-on-task was tracked via the AI interaction platform (Langfuse v3.109.0), which recorded the duration of active engagement (excluding idle time). The experimental group had a mean time-on-task of 41.2 ± 2.3 minutes per session (91.6% of the 45-minute session duration). For the control group, time-on-task was recorded by a trained research assistant using a stopwatch, with a mean time-on-task of 39.8 ± 3.1 minutes per session (88.4% of the session duration). No significant differences in time-on-task were observed between groups ($t(18)=1.21$, $p=0.24$).

Completion Rates: Completion rates were defined as the percentage of assigned intervention tasks completed by each participant. For the experimental group, tasks included completing all AI-simulated ethical scenarios ($n=6$ total scenarios across 8 sessions) and providing meaningful responses to AI feedback. All 10 participants in the experimental group completed 100% of the assigned AI scenarios and feedback responses, resulting in a 100% completion rate. For the control group, tasks included completing weekly ethical discussion prompts ($n=8$ prompts) and reading assigned ethics materials ($n=4$ readings). Nine of 10 control group participants completed 100% of the tasks, with one participant missing one discussion prompt (attributed to a temporary technical issue during an online session), resulting in an overall completion rate of 97.5% for the control group.

Number of Interactions per Participant: For the experimental group, the number of interactions per participant was defined as the total number of distinct exchanges between the participant and the AI system (azure/gpt-4.1-chatbot) during the simulated scenarios. Interactions were tracked via AI interaction logs, with each participant's total interactions calculated as the sum of their responses to AI prompts and follow-up exchanges. The experimental group had a mean of 14.3 ± 2.1 interactions per participant (range: 11–18 interactions). The control group had no AI interactions; their engagement was measured via verbal contributions during group discussions, with a mean of 5.2 ± 1.4 verbal contributions per participant per session (not equivalent to AI interactions, reported for contextual comparison only).

Fidelity Checks (Observer Checklist): Two independent trained research assistants (blinded to group assignment) conducted fidelity checks using a standardized observer checklist for 50% of the sessions (4 sessions total, 2 per group). The checklist included 10 items assessing adherence to the intervention protocol (e.g., “Experimental group participants engaged with AI simulations as instructed,” “Control group discussions focused on assigned ethical topics,” “Intervention leader followed session scripts”). Each item was rated on a 3-point scale (1=not adherent, 2=partially adherent, 3=fully adherent). The mean fidelity score across all checked sessions was 2.9 ± 0.1 (out of 3.0), indicating high adherence to the intervention protocol. No instances of non-adherence were recorded; partial adherence (score=2) was noted in one control group session (one discussion briefly diverted from the assigned topic, corrected within 2 minutes). Inter-observer agreement for the checklist was high (Cohen's $\kappa=0.86$), confirming the reliability of the fidelity checks. Disagreements between observers ($n=2$ minor discrepancies) were resolved via joint discussion and consultation with the study lead.

Appendix C

Data Availability Statement

This study was conducted in strict compliance with the ethical guidelines for human subject research outlined in the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board (IRB) of the Faculty of Educational Studies, Universiti Putra Malaysia prior to the initiation of data collection. All participants provided written informed consent after receiving a detailed explanation of the study's purpose, procedures, potential risks (minimal, related to emotional discomfort from ethical dilemmas), and rights (e.g., voluntary withdrawal without penalty, anonymity of data). To protect participant privacy, all personal identifiers (e.g., names, student IDs, contact information) were removed from the dataset during processing, and participants were assigned unique anonymous codes for data analysis and storage.

The anonymized datasets and analytical code generated during the current study are publicly available in a stable, persistent repository to ensure transparency and reproducibility. (<https://figshare.com/10.6084/m9.figshare.24578913>.)

Predictors of Cross-Cultural Adaptability Among International Postgraduate Students in A Malaysian Research University: A Quantitative Correlational Study

ABSTRACT

Authors: Siti Noormi Alias¹ and Nur Aina Khadijah Azhari ^{1*}

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:
sitinoormi@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Despite the increasing enrolment of international students in Malaysian public universities, empirical understanding of the cross-cultural adaptation process within the Malaysian context remains limited. Grounded in Acculturation Theory, this study examines the predictive roles of emotional intelligence, resilience, and social support to cross-cultural adaptability among international postgraduate students. A quantitative research design was employed, with data collected from international postgraduate students using a random sampling technique. Validated instruments were utilised, including the Trait Emotional Intelligence Questionnaire, Resilience Scale, Perceived Social Support Questionnaire, and the Generic Sense of Ability to Adapt Scale (GSAAS). Multiple regression analysis revealed that emotional intelligence, resilience, and social support collectively explained 33% of the variance in cross-cultural adaptability. Resilience emerged as the strongest predictor ($\beta = .468, p < .001$), followed by emotional intelligence ($\beta = .338, p < .001$) and social support ($\beta = .261, p < .001$). The findings indicate that internal psychological resources, particularly resilience and emotional intelligence, play a dominant role in facilitating effective cultural adjustment, while social support serves as an important contextual enabler. This study contributes to the literature by empirically providing support consistent with Acculturation Theory within the Malaysian higher education context and highlights the need for university policies and intervention programmes that strengthen students' psychological readiness and institutional support mechanisms to enhance successful cross-cultural adaptation.

Keywords: Cross-Cultural Adaptability, Resilience, Social Support, Emotional Intelligence, Malaysian Culture, Higher Education, International Students

1. Introduction

The mobility of international students has become a defining feature of global higher education, reflecting the increasing interconnectedness of academic systems and the intensifying competition among universities to build international visibility. Worldwide, the number of students pursuing education outside their home countries continues to rise, driven by the pursuit of high-quality academic opportunities, research exposure, and enhanced employability in a globalised labour market. International students contribute significantly to the cultural, academic, and economic vitality of host institutions, enriching campus diversity and enhancing cross-cultural learning environments. Yet, despite these contributions, a substantial body of research highlights that international students frequently encounter complex challenges in adjusting to new academic systems, communication norms, and cultural contexts, making cross-cultural adaptability an essential component of their overall student experience (Zhai & Razali, 2022).

Globally, postgraduate enrolment is increasing as universities place greater emphasis on research productivity, knowledge innovation, and advanced professional training. This trend positions postgraduate international students at the centre of global research and intellectual exchange. However, postgraduate studies often require higher levels of academic autonomy, critical thinking, and engagement in scholarly communities, in which demands that may intensify the pressures associated with cultural adjustment (Sam, Jamil, & Md Zain, 2024). Within this international landscape, Malaysia has emerged as a competitive destination, attracting students from more than 150 countries. As of the third quarter of 2025, a total of 23,633 international postgraduate students were registered in Master's and PhD programmes across Malaysian universities, signalling strong growth in the nation's internationalisation agenda (Education Malaysia Global Services, 2025). Malaysia's aspiration to become a prominent regional and global education hub is supported by government initiatives that promote research collaboration, international academic partnerships, and institutional excellence.

Malaysia's unique sociocultural environment further enhances its attractiveness for international students. As a multiracial and multireligious nation, Malaysia embraces cultural pluralism while positioning Islam as an important element of its national identity (Ebrahimi, Yusoff, & Idris, 2025). Its longstanding tradition of religious tolerance and multicultural coexistence has contributed to social stability, making it a compelling environment for international learners seeking a supportive and diverse cultural setting (Zawawi, Zain, Embong, & Alwi, (2022). The unique multicultural environment in Malaysia allows it to be one of the most important centres for international education, thus drawing foreigners into its culturally rich and accommodative environment (Adams, & Velarde, 2021). However, despite of these strengths, the rapid increase in international student numbers has raised concerns related to institutional readiness to provide adequate academic, psychological, and social support among researchers and practitioners including higher education providers (HEIs) (Setti, Sommovigo, & Argentero, 2022). Consequently, there have been numerous calls for more research aimed at identifying factors driving cross-cultural adaptability among international postgraduate students. International students often face challenges such as cultural distance, language barriers, limited social support networks, and heightened academic pressures, all of which may affect their well-being and academic performance (Zhai & Razali, 2022).

These issues highlight several important research gaps. Empirically, most studies on international student adjustment have been conducted in western and developed countries, with limited investigations focusing on Southeast Asia or the Malaysian context. Moreover, existing studies in Malaysia tend to examine general forms of adjustment without addressing the psychological and socio-emotional factors such as resilience, emotional intelligence, and social support that may facilitate cross-cultural adaptability among postgraduate international students. Cross-cultural adaptability briefly describes the extent one is able to adapt and thrive within another culture. Detailed conceptualization of cross-cultural adaptability will be discussed in the subsequent section. Theoretically, there is insufficient integration of more contemporary psychological theories that emphasise adaptive capacities. The present study addresses this gap by drawing on Acculturation Theory which captures psychological and social learning into cultural navigation processes.

Despite Malaysia's internationalisation agenda, there remains a practical gap in institutional strategies that support international postgraduate students. Current support systems are often generic and not tailored to the varying emotional and psychological needs of international learners, especially those enrolled in research-intensive programmes. Identifying the roles of resilience, emotional intelligence, and social support can inform more targeted and evidence-driven interventions that enhance international students' adaptability and overall academic experience. Therefore, this study aims to investigate the predictive roles of resilience, emotional intelligence, and social support to cross-cultural adaptability among international postgraduate students at a selected Malaysian research university. Findings of this study are expected to benefit HEIs, ministries, and non-governmental agencies in navigating social support and cultural exchange programs among postgraduate international students.

This study extends the literature by (a) analysing international postgraduate students at a selected Malaysian research university; (b) using the GSAAS instrument for culture adaptation; and (c) simultaneously modelling internal resources (EI, resilience) and external support (social support) on cross-cultural adaptability.

2. Literature Review

2.1. Conceptualizing Cross-cultural Adaptability

Cross-cultural adaptability refers to the process by which individuals become more proficient and compatible with a new culture. It involves both the willingness and the capacity of individuals to adjust their behaviour in accordance with the dominant norms, values, beliefs, customs, and social expectations that characterize a particular society or geographical context (Chen, Xu, Shi & Liu, 2025). It can be understood as a dynamic process in which individuals, shaped by cultural influences, continuously regulate and adjust themselves to unfamiliar cultural environments, leading to the development of cultural identity (Chen, 2022). Setti, Sommovigo, and Argentero (2022) asserted that psychological comfort in a foreign country can be considered as an outcome of successful cross-cultural. This process involves adjusting to culture shock, achieving psychological adaptation, and interacting effectively within the new cultural context (Chang, Yuan, & Chuang, 2013). The literature on cross-cultural adjustment generally examines it through two primary lenses, which is stress and coping, or culture learning (Young & Schartner, 2014) as the phenomenon took place in the social setting.

International education, in particular, provide valuable opportunities for cross-cultural learning and communication. However, they may also face several challenges including language barrier, diverse educational and cultural backgrounds, relational difficulties, loneliness, and racial discrimination that hinder their ability to fully engage in new learning experiences (Zhai & Razali, 2022). Therefore, international students often struggle to interact with local students and adapt to the host country environment. Overcoming these obstacles is essential for successful adjustment. In accordance, cross-cultural adaptation can be understood as the process through which individuals acquire new cultural knowledge, develop necessary skills, and gradually integrate into the host culture. In this study, cross-cultural adaptability is conceptualised as a general personal capacity to adjust effectively across diverse cultural contexts, consistent with the operationalisation of the Generic Sense of Ability to Adapt Scale (GSAAS) (Franken et al., 2023). In contrast, cross-cultural adjustment/adaptation is treated as a context-specific outcome, reflecting an individual's psychological and behavioural functioning within a particular host culture. Cross-cultural adaptability in this study represents a broader, trait-like capability that facilitates successful adjustment outcomes in specific intercultural settings.

2.2 Mechanism of Cross-cultural Adaptability from Acculturation Theory

Acculturation Theory provides a useful framework for explaining the mechanisms of cross-cultural adaptation and the roles of emotional intelligence, resilience, and social support in facilitating successful adjustment. The theory posits that cross-cultural adaptability involves individuals' psychological and behavioural capacity to adjust and "fit in" with a new cultural environment. According to Berry (1997), acculturation involves four strategies: integration, assimilation, separation, and marginalisation. According to this framework, individuals adopt one of four acculturation strategies based on their orientation towards their original culture and the host culture: integration (maintaining one's original culture while engaging with the host culture), assimilation (relinquishing the original culture in favour of the host culture), separation (maintaining the original culture while avoiding interaction with the host culture), and marginalisation (loss of connection to both original and host cultures).

This adaptive process reflects a natural human tendency to seek internal equilibrium when confronted with unfamiliar social norms, values, and practices. Adaptation often begins with psychological and physiological strain resulting from displacement, commonly recognised as culture shock (Kim, 2017). Through continuous cultural learning, interaction, and experience within the host environment, individuals gradually develop greater emotional stability and functional competence. Over time, these cumulative experiences enable individuals to perform more effectively and attain psychological comfort in the host society. Although integrative communication theories (e.g., Kim, 2017) highlight communication competence in adaptation, this study is grounded in Acculturation Theory, focusing on cultural orientation and acculturation strategies.

Within Acculturation Theory, environmental conditions are recognised as important factors that shape individuals' ability to adopt adaptive acculturation strategies such as integration. In particular, external resources, especially social and institutional support within the host country can facilitate more effective cultural engagement by reducing acculturative stress and promoting psychological well-being (Berry, 2006; Ward & Fischer, 2008). Such support systems act as contextual enablers that interact with individual-level characteristics, including emotional intelligence, in shaping cross-cultural adaptability.

Individuals with higher emotional intelligence are better equipped to perceive, regulate, and respond appropriately to emotional experiences arising from cultural transitions. These internal capabilities are further reinforced by supportive environments, where access to peer networks, institutional guidance, and inclusive policies enhances individuals' confidence and capacity to engage with the host culture. Similarly, resilience supports sustained adaptive functioning in the face of cultural challenges (Fletcher & Sarkar, 2013), particularly when complemented by external support mechanisms.

In the Malaysian higher education context, institutional practices such as the use of English as the medium of instruction, orientation programmes, peer-mentoring systems, counselling services, and culturally inclusive campus initiatives can be understood as structural facilitators that support the integration strategy outlined in acculturation theory. These initiatives reduce communication barriers, increase accessibility to academic and social resources, and promote a sense of belonging among international students. Rather than serving as background context alone, these institutional features function as enabling conditions that interact with individual psychological resources, such as emotional intelligence to influence cross-cultural adaptability outcomes. When both internal capacities and external institutional supports are aligned, international students are more likely to achieve effective adaptation and sustain positive academic and social functioning in the host environment.

2.3. Postgraduate Studies in Malaysia

International education is perceived as important for career development as it provides the individual with international experience (Chang, Yuan, & Chuang, 2013). Singh and Jack (2018) specifically highlight three major benefits of studying abroad that are associated with economic, educational, social, and cultural pull factors, namely enhanced academic achievement, development of knowledge and skills, and meaningful contributions to the home country upon return. Postgraduate students, who are often regarded as adult learners, differ from traditional-aged students in several important ways. They are typically self-directed and autonomous, and they enter higher education voluntarily and with clearer personal and professional goals. Consequently, they are generally expected to take greater responsibility for their learning processes and broader life decisions (Kapur, 2015).

In many countries, including Malaysia, postgraduate education is commonly offered in three modes, namely research, mixed mode and coursework, at both the master's and Doctor of Philosophy (PhD) levels. Research universities are often the preferred choice for students who seek academic specialization and intend to pursue careers in academia, beyond other considerations such as course structure and institutional prestige. In Malaysia, only five universities are designated as research universities, and they are widely recognized for their substantial contributions to research, knowledge production, and academic development. Consequently, it is widely acknowledged that the standards established at research universities are higher than those of other public universities (Altbach, 2007).

The challenges of studying at research universities extend beyond publication requirements and academic performance alone. For international postgraduate students in particular, adapting to Malaysia's cultural context, social norms, and local values can present substantial difficulties if these issues are not addressed proactively. Although English is commonly used as the primary medium of instruction and communication for international students at Malaysian research universities, linguistic accessibility does not necessarily eliminate challenges related to cultural

adjustment, social integration, and everyday interactions. Differences in communication styles, academic conventions, and social practices may continue to affect students' learning experiences and psychological adjustment, thereby influencing both academic outcomes and overall well-being.

However, Malaysia's status as a multiracial society, comprising three major ethnic groups, namely Malays, Chinese, and Indians, may offer certain advantages to international postgraduate students, particularly those who share similarities in customs, food, and cultural practices with one or more of these communities. Such cultural familiarity may facilitate initial adjustment, promote a sense of belonging, and reduce culture shock during the early stages of transition.

2.4. Factors Influencing Cross-cultural Adaptation and Hypothesis Development

Based on the previous discussion, this study highlights the predictive roles of emotional intelligence, resilience, social support and cross-cultural adaptability among international postgraduate students. A review of prior literature suggests that studies integrating all three predictors in the context of cross-cultural adaptability remain scarce. Consequently, there is insufficient evidence to establish the relative dominance of each predictor.

2.4.1. Emotional Intelligence

Emotions are an integral aspect of an individual's personality and life, profoundly influencing personal and professional success. Emotional intelligence, as a cornerstone of human behaviour, enables individuals to effectively express their inner feelings and navigate their environment. By combining emotions with thought, individuals experience a deeper understanding of feelings, which act as powerful motivators, driving immediate action and offering solutions to challenges (Kant, 2019). Conversely, individuals with low emotional intelligence often face difficulties in managing stress, with research underscoring the significant relationship between emotional intelligence and stress management (Sharma & Kumar, 2016). During periods of intense tension, people employ various strategies to cope with stress including search for in depth information and increase frequency of social interactions.

In educational contexts, emotionally supportive environments foster trust and a sense of connection, enhancing students' motivation and academic performance. Such settings encourage risk-taking, exploration of new ideas, and greater satisfaction with studies, contributing to a positive learning climate that booster's academic outcomes (Higgins & BuShell, 2018). Individuals with high emotional intelligence excel in managing interpersonal relationships, regulating their reactions, and controlling negative emotions, thereby maintaining autonomy in their lives (Gebregergis, Huang, & Hong, 2020). This concept hinges on recognizing and understanding one's emotions, guiding emotional responses, and assisting others in managing theirs.

Furthermore, emotional intelligence nurtures self-awareness, empathy, and value-driven behaviours. Students with high emotional intelligence engage in altruistic and assertive actions, participate in team-based activities, and demonstrate leadership qualities. Academic activities such as group projects, role-playing, and collaborative exercises not only build these traits but also complement academic achievements (Chandra, 2021). Research indicates that individuals with higher emotional intelligence are better equipped to manage stress and

navigate the challenges associated with cross-cultural transitions (Shuo et al., 2022). Based on the preceding discussion, despite the absence of specified effect size, existing evidence indicates a potential positive predictive relationship between emotional intelligence and cross-cultural adaptability (e.g., Sundaramoorthy et al., 2025).

This study therefore, hypothesized that:

HA1: There is a significant and positive predictive roles of emotional intelligence on cross-cultural adaptability among international postgraduate students in a Malaysian Research University

2.4.2. Resilience

Resilience, a central concept in positive psychology, is a vital individual resource for coping with and overcoming adversities and perceived stress. It refers to the ability to bounce back (Vella & Pai, 2019). In relations to specific life context, Pooley and Cohen (2010, p. 34) defined resilience as “The potential to exhibit resourcefulness by using available internal and external recourses in response to different contextual and developmental challenges”. Life challenges varies according to individuals. It may be related to the capacity to adjust effectively when confronted with adversity, trauma, significant stress, or life disruptions, including challenges related to personal relationships, serious health conditions, or financial hardship.

Increasingly, research has focused on resilience’s role in mitigating the adverse effects of stressful experiences on psychological outcomes (Chi et al., 2016). Defined as the capacity to recover from adversity or difficult situations (Masten, 2001), resilience has been identified as a critical factor for the successful adjustment of international students. Traits associated with resilience are essential for managing change and are among the most reliable indicators of how well international students adapt to new environments.

Empirical studies reveal that resilience closely related to optimistic characteristics. Optimistic individuals are more likely to exhibiting greater adaptability to challenges than their less optimistic counterparts (Gómez Molinero et al., 2018). Resilience is also vital for maintaining positive mental health, defined as the ability to recover from and effectively manage challenging life situations. Among university and college students, promoting resilience has been shown to alleviate the negative emotions associated with stress and foster academic success (Radhamani & Kalaivani, 2021). Additionally, enhancing resilience has been linked to reduced depression and improved mental health outcomes, including diminished symptoms of anxiety and stress, in both students and adolescents. Based on the above discussion, although the effect size is not specified, there is evidence suggesting a significant predictive influence of resilience and cross-cultural adaptability (e.g., Duanaeva et al., 2023).

Accordingly, this study hypothesized that:

HA2: There is a significant and positive predictive role of resilience and cross-cultural adaptability among international postgraduate students in a Malaysian Research University.

2.4.3. Social Support

Social support refers to the assistance an individual receives through connections with others,

including family, friends, groups, and the broader community. In this study context, it is specified to the host country, which is Malaysia. This support system plays a vital role in individuals' mental health and overall quality of life, particularly among international postgraduate students. The absence of social support has been identified as a significant contributor to mental health challenges, such as depressive symptoms (Bukhari & Afzal, 2017), and negatively impacts students' quality of life (Dafaalla et al., 2016). Research highlights that social support from family and host country friends significantly enhances emotional, social, and academic performance.

Notably, for postgraduate international students, academician and friends, often become a more critical source of social support than family. Identifying specific sources of social support is crucial for promoting mental health and fostering the emotional, social, and academic adjustment of international students (Alsubaie, Stain, Webster, & Wadman, 2019). In particular, perceived support from family and teachers reduces the likelihood of depressive and anxiety symptoms and enhances subjective well-being (Hellfeldt, López-Romero, & Andershed, 2020). Lecturer, in particular, may serve as a critical support system for students experiencing social struggles.

Social support plays a pivotal role in the emotional and psychological adjustment of international students. It serves as a buffer against stress and enhances adaptability, enabling students to cope with the demands of studying abroad (Hidalgo-Fuentes et al., 2024). Moreover, higher perceived social support correlates positively with cultural adjustment, in which it reduces homesickness and academic pressures by reducing stress levels through effective coping strategies. Furthermore, students who actively engage with their social networks through friendships, peer support groups, or institutional resources tend to exhibit better cross-cultural adaptability (Bender et al., 2019). The type of social support received can vary in effectiveness based on cultural contexts. For instance, perceived social support enhances self-evaluations and positive self-cognition, which are critical for adapting to stressful academic environments (Zhou & Xiang, 2025). This suggests that social support play a significant role in individual adaptation processes to new environment. Based on the foregoing discussion, while the effect size remains unspecified, it is evident that social support is likely to have a positive predictive influence on cross-cultural adaptability (e.g., San & Guo, 2023).

Therefore, it is hypothesized that,

HA3: There is a significant and positive predictive roles of social support and cross-cultural among international postgraduate students in a Malaysian Research University.

2.4.4. Research Framework

Figure 1 presents the research framework of the study. The model proposes that emotional intelligence, resilience, and social support significantly enhance the likelihood of successful cross-cultural adaptation among international postgraduate students at a selected research university in Malaysia.

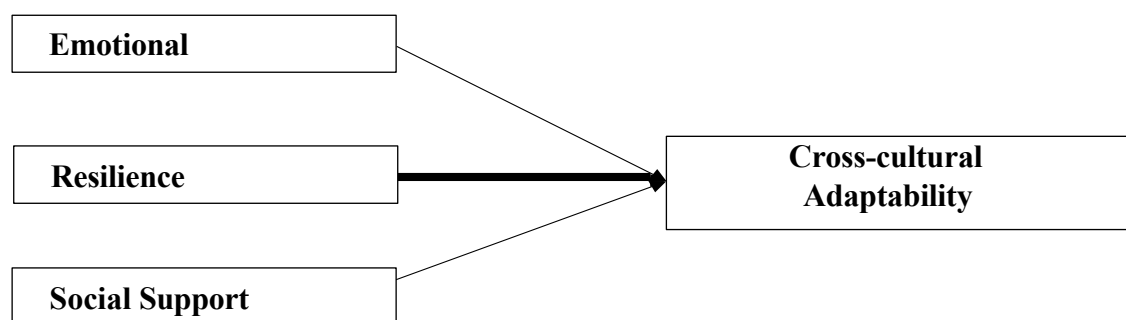


Figure 1. Research Framework

3. Methodology

This study adopts a quantitative, correlational research design to determine the predictive roles of EI, resilience, and social support on cross-cultural adaptability among international postgraduate students at a selected research university (RU) in Malaysia. To date, the RU reported a total of 7,922 international postgraduate students. Raosoft online sample size calculator proposed a minimum of 262 respondents needed for this study. Oversampling was then applied avoid issues of missing or unusable questionnaire resulted in final distributed questionnaire of 328 ($262/.80 = 327.5 \approx 328$). The sample has been selected through simple random sampling method using Microsoft Excel.

This study adopts a quantitative, correlational research design to determine the predictive roles of EI, resilience, and social support on cross-cultural adaptability among international postgraduate students at a selected research university (RU) in Malaysia. To date, the RU reported a total of 7,922 international postgraduate students. Raosoft online sample size calculator proposed a minimum of 262 respondents needed for this study. Oversampling was then applied avoid issues of missing or unusable questionnaire resulted in final distributed questionnaire of 328 ($262/.80 = 327.5 \approx 328$). The sample has been selected through simple random sampling method using microsoft excel.

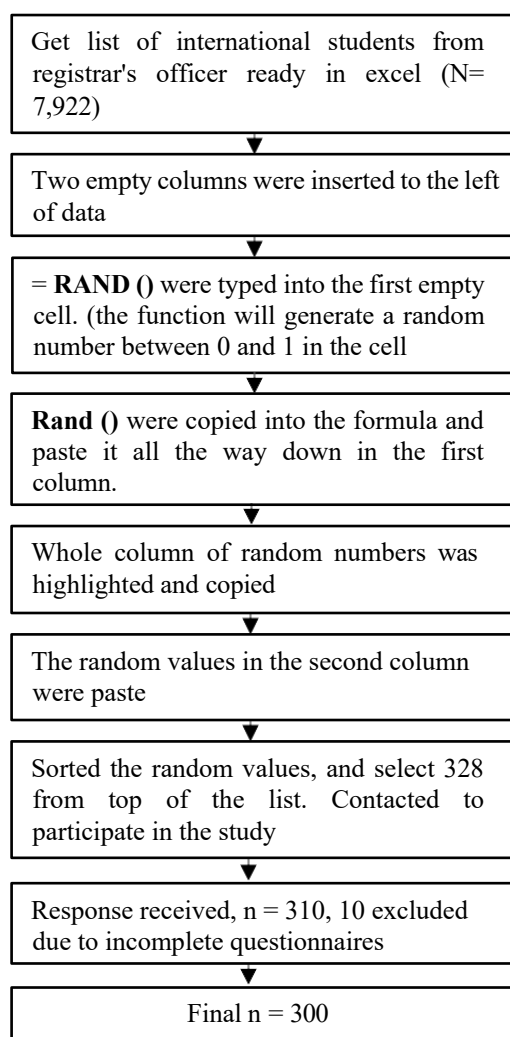


Figure 2. Sampling Procedure

3.1 Measurement and Instrument

A structured questionnaire has been developed based on adaptation made from existing well-established instruments. The questionnaire consists of five sections described as follows. It was distributed to the respondents through Google Form.

3.1.1. Section A: Demographic Information

This section includes the gender, age, year of study and nationality of the respondents.

3.1.2. Section B: Cross-Cultural Adaptability

Cross-cultural adaptability in this study is operationalized as the individual international postgraduate student's capabilities to adjust and adapt to Malaysia cultural environment. This includes understanding and integrating into the social norms, values, and behaviours of the Malaysia's culture while maintaining one's own cultural identity. This study has adapted Generic Sense of Ability to Adapt Scale (GSAAS) developed by Franken et al. (2023) to

measure cross-cultural adaptability. The instrument consists of 10 items and measured based on five- likert scale ranging from 1 = Strongly Disagree to 5 = Strongly Agree. Based on past studies, the reliability value of the instrument is 0.871. Example of an item is "I can cope well with my daily life in Malaysia".

GSAAS is a relatively new instrument that has been previously validated within a mental health outpatient sample. While the scale has demonstrated acceptable reliability and validity in its initial context, its application in the current study extends to a different population, namely international university students. To ensure its appropriateness for this context, the instrument underwent a validation process involving two experts review (discussed in the next section), where subject matter experts evaluated the items for clarity, relevance, and contextual suitability for the target population. This process helped to establish content validity and ensured that the items were understandable and appropriate for use among international students.

3.1.3. Section C: Emotional Intelligent

Emotional Intelligence in this study is operationalized as the ability to recognize, understand, manage, and influence one's own emotions as well as the emotions of others. This study uses the Trait Emotional Intelligence Questionnaire (TEIQue) developed by Petrides (2009) to measure EI. The instrument consists of 30 items and is measured based on a seven-point Likert scale, ranging from 1 = Completely Strongly Disagree to 7 = Completely Strongly Agree. Cronbach's alpha of the instrument is 0.93 (Petrides et al., 2007), indicating strong internal consistency. Example of an item is "I can deal effectively with Malaysian people." All negative items were reverse coded before the actual analysis conducted.

The TEIQue employs a 7-point Likert scale to capture greater variability and sensitivity in respondents' emotional self-perceptions. In contrast, the other constructs in this study were measured using 5-point Likert scales, which are commonly used in social science research due to their simplicity and ease of response, particularly in multicultural samples. The use of different scale formats reflects adherence to the original validated instruments, thereby preserving their psychometric integrity. As the analyses were conducted using mean scores for each construct, differences in scale range were not expected to substantially bias the results. Nonetheless, scale reliability and variance were carefully examined to ensure comparability across measures.

3.1.4. Section D: Resilience

Resilience in this study is operationalized as the capacity to recover from adversity, adapt to challenges, and maintain psychological well-being in the face of stress. Resilience Scale (Wagnild & Young, 1993) and Workplace Resilience Instrument (Mallak & Yildiz, 2016) were adapted in this study to measure resilience. The instrument consists of 22 items and is measured using a five-point Likert scale, ranging from 1 = Strongly Disagree to 5 = Strongly Agree. The scale has demonstrated acceptable reliability in previous studies, with a Cronbach's alpha value of 0.78, indicating high internal consistency (Wagnild & Young, 1993). Example of an item is "I do not have to rely much on my friends at this university".

3.1.5 Section E: Social Support

Social support in this study is operationalized as the availability of emotional, informational, and practical assistance provided by others, which helps individuals cope with stress. The Perceived Social Support Questionnaire (F-SozU K-14) developed by Kliem et al. (2015), is used to measure social support. The instrument consists of 14 items and is measured using a five-point Likert scale, ranging from 1 = Does not apply to 5 = Exactly applicable. Previous studies have reported good reliability for the F-SozU K-14, with a Cronbach's alpha of 0.94, indicating strong internal consistency (Kliem et al., 2015). Example of an item is "If I need to, I can borrow something from Malaysian friends or neighbours without any problems".

3.2. Validity and Reliability

Validity of the study instrument was determined based on face and content validity. Face validity was conducted by researchers when searching for the most suitable instrument from literature. The researchers have compared the individual items for each instrument with the operational definition of each construct. The researchers have selected the instrument which she perceived "on its surface" as mostly measuring the construct. The principal researcher of this study has more than 10 years conducting research in education and human resource development fields. Meanwhile, content validity of the study has been conducted by two experts which have more than five years working experience in academia and has involves in various types of research related to social adjustments. The experts were selected based on their familiarity with the constructs of emotional intelligence, social support, resilience, and cross-cultural adaptability, as well as their experience in instrument development and validation. They were asked to evaluate each item in terms of its relevance, clarity, and appropriateness for the target population of international students. To start with, the researcher has provided the experts with an excel sheet consisting of conceptual and operational definition of terms as well as list of items for each construct to be reviewed. Feedback and recommendations provided by expert were deeply considered to improve the items.

Next, researcher proceed with reliability analysis, specifically internal reliability using data from pilot and actual studies. Cronbach's alpha values were computed and compared between the pilot and main study phases to assess internal consistency. During the data screening process for the main study, items 7, 10, and 28 from the emotional intelligence construct were removed due to low corrected item-total correlations, as indicated by the "Cronbach's alpha if item deleted" analysis. In addition, items 3, 12, and 23 were excluded due to factor loadings below the acceptable threshold of .50, as presented in Table 3. Following the removal of these items, the emotional intelligence construct was reanalysed, resulting in a final scale comprising 24 items with improved reliability. A coefficient of 0.7 or more is widely viewed as an acceptable level for most social science research studies. Table 1 indicates the Cronbach's Alpha values of all study construct. Accordingly all instrument recorded good reliability.

Table 1. Cronbach's Alpha Values

Construct	No. of Item	Pilot Study (n=25)	No. of Item	Actual Study (n=300)
Cross-cultural Adaptability	10	0.841	10	0.852
Emotional Intelligence	30	0.708	24	0.720
Resilience	22	0.867	22	0.876
Social Support	14	0.807	14	0.861

3.3. Pilot Study

A pilot study has been conducted prior to actual study involving 25 international postgraduate students, which took place around two weeks starting from 02 June 2025 to 16 June 2025. While a sample size of 25 is generally considered acceptable for preliminary reliability testing in social science research, it is acknowledged that the relatively small sample size, particularly in relation to the total number of items (76 items), may limit the precision and stability of the Cronbach's alpha estimates. Therefore, the pilot results should be interpreted as indicative rather than definitive, serving primarily to identify potential issues in item clarity and internal consistency. Respondents were mostly recruited at the university's library. Feedback received by respondents during the pilot have been considered to improve the items structure to be clearer. Those participated during pilot study were eliminated from the population list for actual study.

3.4. Data Collection Procedure

To start with, the researcher managed to obtain approval and list of international postgraduate students from the university's School of Graduate Studies office. 328 sample has been identified from simple random sampling technique using Microsoft Excel. Researchers personally contacted the respondents through email and phone call to make an appointment. After three months, there were 310 respondents completed the study. Upon initial screening, 10 instruments have to be eliminated due to incomplete questionnaire. Respond rate for the study therefore is 95%.

Table 2 indicates the demographic findings of the study. Descriptive statistics were used to summarize the demographic characteristics of the respondents, such as gender, age, year of study, and faculty and institute of study. The data are presented in Table 2 for clarity.

Table 2. Demographic Profile (n = 300)

Demographic Factors	<i>f</i>	%
Gender		
Female	228	76
Male	72	24
Age (years)		
21-25	90	30
26-29	144	48
> 30	66	22
Year of Study (years)		
1	72	24
2	154	51.33
> 3	74	24.67
Nationality		
China	170	56.77
Bangladesh	45	14.99
Indonesia	34	11.48
India	19	6.31
Pakistan	17	5.61

Sudan	15	4.84
Myanmar	14	4.79

The data analysis on gender showed that from 300 respondents, 228 respondents are female (76.0%), aged 144 respondents between 26-29 years old, and at their second year at their second year in the university. On top of that, majority of postgraduate students at the research university are from China (56.77%).

3.5. Data Analysis

The statistical data from this study has been analysed using IBM SPSS version 29.0. Data screening were conducted prior to actual data analysis.

3.5.1. Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) was conducted using principal components with promax rotation. Results in Table 3 indicated that items 3, 12, and 23 under emotional intelligence are recommended to be deleted due to low factor loadings ($<.5$). The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was 0.92, indicating excellent suitability for factor analysis. Bartlett's Test of Sphericity was significant, $\chi^2(1,820) = 19,864.37$, $p < 0.001$, confirming that the correlation matrix was factorable. The extracted factors explained approximately 74% of the total variance, exceeding the recommended threshold for social science research. All retained items demonstrated strong factor loadings and acceptable communalities (≥ 0.30). In addition, based on Common Method Bias (Harman's Single Factor Test) Analysis, the percentage of variance is 32.67%, which is below the recommended threshold of 50%, suggesting that common method bias did not significantly affect the study results.

Table 3. Factor Loadings and Communalities Values

Construct	Items	Factor loading								Communality
		1	2	3	4	5	6	7	8	
Cross-cultural adaptability	8	.949								.962
	2	.932								.885
	1	.932								.882
	10	.801								.871
	6	.960								.909
	4	.924								.755
	5	.924								.758
	3	.573								.882
	7	.573								.962
Emotional Intelligence	9	.529								.885
	2		.930							.865
	4		.916							.839
	5		.966							.933
	6		.684							.468
	9		.953							.908
	15		.506							.526
	16		.868							.754
	18		.868							.754
	30		.556							.539
	8			.979						.958

Construct	Items	Factor loading								Communality
		1	2	3	4	5	6	7	8	
	11			.861						.741
	12			.459						.611
	13			.981						.962
	14			.838						.702
	17			.858						.736
	22			.858						.736
	23			.450						.603
	25			.908						.825
	19				.755					.570
	20				.789					.623
	26				.888					.789
	27				.966					.933
	1					.813				.661
	3					.457				.609
	21					.735				.540
	24					.980				.960
	29					.536				.677
Resilience	1						.719			.725
	2						.782			.785
	3						.851			.754
	4						.791			.818
	5						.777			.695
	6						.815			.669
	7						.913			.587
	8						.882			.624
	9						.782			.795
	10						.913			.607
	11						.851			.764
	12						.791			.828
	13						.808			.687
	14						.814			.732
	15						.778			.552
	16						.828			.808
	17						.851			.774
	18						.823			.842
	19						.722			.743
	20						.828			.838
	21						.814			.722
	22						.851			.724
Social Support	1							.772		.635
	3							.931		.965
	4							.915		.869
	5							.915		.988
	6							.736		.988
	8							.831		.820
	9							.915		.965
	10							.759		.739
	11							.701		.988
	12							.915		.663
	13							.857		.868
	14							.839		.988
	2								.746	.846
	7								.745	.778

Data shows in Table 8 (based on correlation table) indicated that there are no issues on

multicollinearity between variables. Based on collinearity statistics in Table 9, Data conformed with all the assumptions needed for multiple linear regression analysis (Tolerance values > .1, VIF <10). Researcher then proceed with descriptive, correlation, and regression analysis.

3.5.2. Descriptive Analysis

Descriptive analysis (e.g., frequency, percentage, mean, and standard deviation) was conducted to analysed demographic information and to determine the level of constructs.

Table 4. Group Interval for 7-point Likert Scale

Group Interval	Level Description
1.00 – 3.11	Low (1)
3.12 – 5.03	Moderate (2)
5.04 – 7.00	High (3)

Source: Pallant (2020)

Table 5. Group Interval for 5-point Likert Scale

Group Interval	Level Description
1.00 – 2.330	Low (1)
2.331 – 3.660	Moderate (2)
3.661 – 5.00	High (3)

Source: Pallant (2020)

To determine the level of variables, the scales indicated in Table 4 and 5 were referred to.

3.5.3. Inferential Analysis

Pearson product moment correlation and multiple linear regression were used to determine the relationships between resilience, emotional intelligence, social support, and cross-cultural adaptability international postgraduate students at a selected research university in Malaysia. Hypothesis testing specifically based on the correlation analysis Guilford's rule of Thumb indicated in Table 6 is used to interpret the nature of relationship between variables.

Table 6. Guilford's Rule of Thumb

r value < 0.20	Correlation Strength
< .20	Negligible
.20 - .40	Low
.40 - < .70	Moderate
0.70 - < 0.90	Strong
0.90 - 1.00	Very Strong

4. Findings and Discussion

Findings of the study has been arranged according to statistical level of analysis begin with the descriptive analysis on the level of variables, correlation, and followed by regression analysis.

Table 7. Descriptive Analysis

Variable	Frequency	Percentage	Mean	SD
Emotional Intelligence	0	0	4.61	0.33
Low (1 – 3.11)	260	86.7		
Moderate (3.12 – 5.03)	60	13.3		
High (5.04 – 7.00)				
Resilience	0	0	4.22	0.59
Low (1 – 2.330)	56	18.7		
Moderate (2.331 – 3.660)	244	81.3		
High (3.661 – 5.00)				
Social Support	0	0	4.22	0.59
Low (1 – 2.330)	56	18.7		
Moderate (2.331 – 3.660)	244	81.3		
High (3.661 – 5.00)				
Cross-cultural Adaptability	0	0	4.25	0.62
Low (1 – 2.330)	56	18.7		
Moderate (2.331 – 3.660)	244	81.3		
High (3.661 – 5.00)				

Table 7 indicated that there is a high level of EI at ($M = 4.61$, $SD = 0.33$), resilience ($M = 4.22$, $SD = 0.59$), social support was ($M = 4.22$, $SD = 0.59$), and cross-cultural adaptability ($M = 4.25$, $SD = 0.62$) among respondents of this study. Therefore, it can be interpreted that all of these international postgraduate students, which considered as adult learners are mentally (emotional intelligence and resilience) prepared with their decision to study abroad.

Next, Table 8 highlights the correlation analysis results between cross-cultural adaptability (the dependent variable) and the independent variables. Emotional Intelligence (EI), Resilience, and Social Support. The analysis was conducted using Pearson's correlation coefficient.

Table 8. Correlation Matrix between Constructs (n=300)

Construct	Cross- cultural Adaptability	EI	Resilient
Emotional intelligence	326**		
Resilience	748**	0.390**	
Social support	547**	0.307**	0.281**

Note: ** result is significant at 0.01 level (2-tailed).

Results in Table 8 show that emotional intelligence ($r = 0.326$) has positive and low relationship with cross-cultural adaptability. Resilience ($r = 0.748$) indicated positive and strong relationship with cross-cultural adaptability. Meanwhile, social support ($r = 0.547$) was found to relate moderately and positively to cross-cultural adaptability. Further reinforcement of this

conclusion is provided by the regression analyses reported in Table 9, which confirm the significant predictive effects of these variables on cross-cultural adaptability.

Prior to conducting multiple linear regression analysis, several key assumptions were assessed to ensure the validity of the results. Normality of residuals was examined through visual inspection of histograms and normal probability (P-P) plots, which indicated that the residuals were approximately normally distributed. Linearity and homoscedasticity were assessed through scatterplots of standardized residuals against predicted values, which showed no clear patterns, suggesting that the assumptions of linearity and homoscedasticity were met. Independence of errors was evaluated using the Durbin–Watson statistic, with values falling within the acceptable range, indicating no significant autocorrelation. Multicollinearity was assessed using Variance Inflation Factor (VIF) values, all of which were below the recommended threshold, confirming the absence of multicollinearity issues among the predictor variables.

Table 9. Regression Analysis between EI, Resilience and Social Support towards Cross-cultural Adaptability

Variable	Unstandardized Coefficients B	Std. Error	Standardized Coefficients, Beta	<i>t</i>	Sig - <i>p</i>	95% confidence interval for B		Collinearity statistics	
						Lower bound	Upper bound	Tolerance	VIF
Constant	1.867	0.141	-	3.271	0.000	2.144	1.590	-	-
Emotional intelligence	.303	0.059	.338	3.193	0.000	.086	1.020	.62	1.61
Resilience	.624	0.210	.468	3.475	0.000	.060	1.478	.55	1.82
Social Support	.237	0.230	.261	2.410	0.000	.084	1.081	.68	1.47

$R = 0.571$; $R^2 = 0.326$; Adj. $R^2 = 0.325$; $F = 6.311$, Sig- $F = 0.000^b$

The multiple linear regression analysis indicated that the overall model was statistically significant ($F = 6.311$, $p < 0.001$), demonstrating a good fit to the data. The R^2 value of 0.326 suggests that approximately 33 per cent of the variance in cross-cultural adaptability is explained by the predictors. Furthermore, the adjusted R^2 of 0.325 indicates that the model maintains reasonable explanatory power. A detailed examination of the standardized coefficients revealed that resilience was the strongest predictor of cross-cultural adaptability among international postgraduate students ($\beta = 0.468$, $p < 0.001$). This was followed by emotional intelligence ($\beta = 0.338$, $p < 0.001$) and social support ($\beta = 0.261$, $p < 0.001$), both of which also contributed significantly to the model. Therefore, it can be concluded that the results of this study support HA1, HA2, and HA3.

Results of this study conform with Acculturation Theory in which cross-cultural adaptations involves internal psychological process (emotional intelligence and resilience) and interactions with external environment (social support). From an acculturation theory perspective, resilience refers to an individual's capacity to adapt, recover, and maintain psychological well-being when confronted with cultural transitions and stressors associated with living in a new sociocultural environment. In the context of international postgraduate students, resilience is not merely the ability to "bounce back" from difficulties, but a dynamic process of ongoing adjustment that enables individuals to manage culture shock, language barriers, academic pressure, and social isolation while continuing to function effectively in the host culture (Vella & Pai, 2019).

A significant relationship between emotional intelligence and cross-cultural adaptability is both theoretically grounded and empirically supported in the literature on intercultural adjustment and acculturation. From an acculturation perspective, cross-cultural adaptation involves continuous psychological and sociocultural adjustment when individuals encounter unfamiliar norms, values, and interactional styles. Individuals with higher levels of emotional intelligence are better equipped to interpret emotional cues across cultures, manage uncertainty, and regulate stress associated with culture shock. This emotional competence enables them to respond more constructively to intercultural difficulties such as miscommunication, social exclusion, and academic pressure, thereby facilitating smoother adaptation. Empirically, studies consistently show that emotionally intelligent individuals demonstrate superior cross-cultural adjustment, particularly in emotional and social domains. For instance, Gebregergis, Huang, and Hong (2020) found that international students with higher emotional intelligence reported fewer adjustment difficulties when dealing with culturally diverse environments.

A significant relationship between social support and cross-cultural adaptability is well established in acculturation and intercultural adjustment literature, particularly in explaining how individuals cope with the psychological and social demands of living in a new cultural environment. Social support refers to emotional, informational, and instrumental assistance received from social networks such as peers, family, academic staff, and the host community. Within the framework of acculturation theory, social support functions as a critical external resource that facilitates both psychological well-being and sociocultural competence during cultural transition. Social support acts as a stress buffer by providing reassurance, guidance, and a sense of belonging, which reduces feelings of isolation, anxiety, and uncertainty. International students who perceive stronger social support networks are more likely to cope effectively with cultural challenges, maintain motivation, and demonstrate positive adjustment outcomes in the host environment (Hellfeldt, López-Romero, & Andershed, 2020).

5. Conclusions and Implications

This study examined the levels and relationships among emotional intelligence, resilience, social support, and cross-cultural adaptability among international postgraduate students at a selected research university in Malaysia. The findings indicate that the overall levels of emotional intelligence, resilience, social support, and cross-cultural adaptability were high, suggesting that the participants generally possessed strong internal and external resources to manage the demands of studying in a culturally different environment. More importantly, all three variables were found to have significant positive relationships with cross-cultural adaptability, confirming their critical roles in facilitating successful cultural adjustment.

Among the predictors, resilience emerged as the most dominant factor influencing cross-cultural adaptability, followed by emotional intelligence and social support. This pattern suggests that personal coping capacity and psychological strength are more influential than external resources in determining students' ability to function effectively in a new cultural context. While emotional intelligence contributes to adaptability through emotional regulation and interpersonal competence, social support plays a complementary role by providing emotional reassurance and practical assistance. Collectively, the findings reinforce Acculturation Theory, which posits that successful adaptation results from the interaction between individual psychological resources and environmental support systems.

The results extend Acculturation Theory by empirically demonstrating the combined influence of psychological traits and social resources on cross-cultural adaptability. The dominant role of resilience highlights the importance of viewing adaptation as a dynamic coping process rather than a purely cultural learning outcome. This study also strengthens the conceptual integration between resilience theory and emotional intelligence frameworks within the intercultural adjustment literature, suggesting that internal psychological capacities are central mechanisms underlying successful adaptation.

For higher education institutions, the findings underscore the need to prioritise psychological preparedness in addition to academic readiness for international students. Universities should consider integrating resilience-building programmes, emotional intelligence training, and coping skills workshops into student development initiatives, orientation programmes, and counselling services. Such interventions may include stress management training, emotional awareness exercises, and reflective practices to strengthen students' adaptive capacity.

Moreover, although social support was not the strongest predictor, its significant role indicates the necessity of maintaining a supportive campus environment. Institutions should strengthen peer mentoring systems, buddy programmes, intercultural activities, and academic advising services to promote social connectedness and reduce feelings of isolation among international students. Academic staff and administrative personnel should also be trained to adopt culturally responsive approaches in teaching and student support.

This study has several limitations that should be acknowledged. First, the sample composition shows an overrepresentation of Chinese nationals (56.77%) and a relatively smaller proportion of participants from other nationalities. While this distribution may reflect the accessibility of respondents within the study context, it may not fully represent the broader population of international students from more diverse cultural backgrounds. As such, the findings should be interpreted with caution, particularly when generalising to non-Chinese international student groups. Second, the gender distribution of the sample is skewed toward female respondents (76%), which may limit the generalisability of the results across gender groups. Differences in experiences, coping styles, and social support mechanisms between male and female students may potentially influence cross-cultural adaptability, but this study did not examine gender as a moderating variable.

Future research is encouraged to investigate whether the relationships among emotional intelligence, social support, resilience, and cross-cultural adaptability differ across gender and nationality groups. Next, considering GSAAS as a relatively new instrument that was originally validated within a mental health outpatient sample. Even though expert validation supports the suitability of the scale for the present sample, differences in population characteristics may still influence how respondents interpret and engage with the items. Therefore, caution is warranted when generalising the findings, and future research is encouraged to further examine the psychometric properties of the GSAAS across diverse non-clinical populations to strengthen its cross-context applicability.

Future studies are recommended to adopt longitudinal designs to examine changes in emotional intelligence, resilience, and social support across different stages of students' academic journeys. Such designs would provide deeper insights into how adaptability develops over time and how students respond to evolving academic and cultural demands. Further research should also explore potential mediating and moderating variables, such as coping strategies,

motivation, personality traits, cultural intelligence, or institutional support, to better understand the mechanisms through which resilience and emotional intelligence influence cross-cultural adaptability. Qualitative or mixed-method approaches are likewise encouraged to capture students' lived experiences and provide richer contextual understanding of their adaptation processes.

Additionally, comparative studies across multiple institutions or countries would be valuable in determining whether the dominance of resilience is culturally or contextually specific. This would enhance the generalisability of the findings and contribute to cross-national perspectives on international student adjustment. Finally, future research could evaluate the effectiveness of institutional interventions by examining the impact of resilience training, emotional intelligence workshops, and social integration programmes on international students' academic outcomes and psychological well-being. Such studies would provide evidence-based guidance for developing targeted support systems in higher education institutions.

Ethics Approval

This study was reviewed and approved by the Universiti Putra Malaysia Ethics Committee (Approval No.: JKEUPM-2023-277). All procedures performed in this study involving human participants were conducted in accordance with the ethical standards of the institutional research committee and the 1964 Helsinki Declaration and its later amendments.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request. Data are not publicly available due to ethical restrictions and participant confidentiality agreements.

Declaration on the Use of Artificial Intelligence

The authors declare that generative artificial intelligence tools were used solely for language editing and grammar checking purposes during the preparation of this manuscript. No artificial intelligence tools were used to generate research data, perform data analysis, or interpret results. The authors take full responsibility for the content of the manuscript.

References

- Adams, D., & Velarde, J. M. (2021). Leadership in a culturally diverse environment: perspectives from international school leaders in Malaysia. *Asia Pacific Journal of Education*, 41(2), 323-335.
- Alsubaie, M. M., Stain, H. J., Webster, L. A. D., & Wadman, R. (2019). The role of sources of social support on depression and quality of life for university students. *International journal of adolescence and youth*, 24(4), 484-496.
- Altbach, P. G. (2007). Peripheries and centres: Research universities in developing countries. *Higher education management and policy*, 19(2), 111.
- Bender, M., van Osch, Y., Slegers, W., & Ye, M. (2019). Social support benefits psychological adjustment of international students: Evidence from a meta-analysis. *Journal of Cross-Cultural Psychology*, 50(7), 827-847.
- Berry, J. W. (2006). Acculturation: A conceptual overview. *Acculturation and parent-child relationships*, 13-32.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology: An International Review*, 46(1), 5-34.
- Bukhari, S. R., & Afzal, F. (2017). Perceived social support predicts psychological problems among university students. *The International Journal of Indian Psychology*, 4(2), 18-27.
- Chang, W. W., Yuan, Y. H., & Chuang, Y. T. (2013). The relationship between international experience and cross-cultural adaptability. *International Journal of Intercultural Relations*, 37(2), 268-273.
- Chandra, Y. (2021). Online education during COVID-19: perception of academic stress and emotional intelligence coping strategies among college students. *Asian education and development studies*, 10(2), 229-238.
- Chen, X. (2022). Cross-cultural competence and workplace adaptability in the perspective of globalization. *Journal of Sociology and Ethnology*, 4(9), 88-93.
- Chen, W., Xu, W., Shi, Q., & Liu, Z. (2025). Factors influencing cross-cultural adaptation and life satisfaction in multilingual contexts: the mediating role of cultural intelligence. *Journal of Multilingual and Multicultural Development*, 1-19.
- Chi, P., Li, X., Du, H., Tam, C. C., Zhao, J., & Zhao, G. (2016). Does stigmatization wear down resilience? A longitudinal study among children affected by parental HIV. *Personality and Individual Differences*, 96, 159-163. <https://doi.org/10.1016/j.paid.2016.03.001>
- Dafaalla, M., Farah, A., Bashir, S., Khalil, A., Abdulhamid, R., Mokhtar, M., ... & Abdalrahman, I. (2016). Depression, anxiety, and stress in Sudanese medical students: a cross sectional study on role of quality of life and social support. *Am J Educ Res*, 4(13), 937-942.
- Duanaeva, S., Berdibayeva, S., Garber, A., Baizhumanova, B., & Adilova, E. (2023). Cross-cultural study of resilience, stress, and coping behavior as prerequisites for the success of international students. *The Open Psychology Journal*, 16(1), 1-13.
- Ebrahimi, M., Yusoff, K., & Idris, A. (2025). Unity in diverse society of Malaysia. *Journal of Lifestyle and SDGs Review*, 5(1), e03493-e03493.
- Education Malaysia Global Services (2025). Student Data: International Student Data 2025. <https://educationmalaysia.gov.my/more/student-data/international-student-data>. Accessed on 27 November 2025
- Fletcher, D., & Sarkar, M. (2013). Psychological resilience. *European psychologist*.
- Franken, K., Schuffelen, P., Ten Klooster, P., van Doesum, K., Westerhof, G., & Bohlmeijer, E. (2023). Introduction of the generic sense of ability to adapt scale and validation in a sample of outpatient adults with mental health problems. *Frontiers in psychology*, 14, 985408.

- Gebregergis, W. T., Huang, F., & Hong, J. (2020). The impact of emotional intelligence on depression among international students studying in China: The mediating effect of acculturative stress. *International Journal of Intercultural Relations*, 79, 82-93.
- Gómez Molinero, R., Zayas García, A., Ruiz González, P., & Guil, R. (2018). Optimism and resilience among university students.
- Hellfeldt, K., López-Romero, L., & Andershed, H. (2020). Cyberbullying and psychological well-being in young adolescence: the potential protective mediation effects of social support from family, friends, and teachers. *International journal of environmental research and public health*, 17(1), 45.
- Hidalgo-Fuentes, S., Martínez-Álvarez, I., Sospedra-Baeza, M. J., Martí-Vilar, M., Merino-Soto, C., & Toledano-Toledano, F. (2024, March). Emotional intelligence and perceived social support: Its relationship with subjective well-being. In *Healthcare* (Vol. 12, No. 6, p. 634). MDPI.
- Higgins, K., & BuShell, S. (2018). The effects on the student-teacher relationship in a one-to-one technology classroom. *Education and Information Technologies*, 23(3), 1069-1089.
- Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist*, 18(1), 12-23.
- Kant, R. (2019). Emotional intelligence: A study on university students. *Journal of Education and Learning (EduLearn)*, 13(4), 441-446. <https://doi.org/10.11591/edulearn.v13i4.13592>
- Kapur, S. (2015). Understanding the characteristics of an adult learner. *Jamia Journal of Education*, 2(1), 111-121.
- Kim, Y. (2017). Integrative communication theory of cross-cultural adaptation., 1-13. <https://doi.org/10.1002/9781118783665.ieicc0041>
- Kliem, S., Möble, T., Rehbein, F., Hellmann, D. F., Zenger, M., & Brähler, E. (2015). A brief form of the Perceived Social Support Questionnaire (F-SozU) was developed, validated, and standardized. *Journal of clinical epidemiology*, 68(5), 551-562.
- Mallak, L. A., & Yildiz, M. (2016). Developing a workplace resilience instrument. *Work*, 54(2), 241-253.
- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American psychologist*, 56(3), 227.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). McGraw-Hill Education.
- Petrides, K. V. (2009). Psychometric properties of the trait emotional intelligence questionnaire (TEIQue). In *Assessing emotional intelligence: Theory, research, and applications* (pp. 85-101). Boston, MA: Springer US.
- Pooley, J. A., & Cohen, L. (2010). Resilience: A definition in context. *Australian Community Psychologist*, 22(1), 30-37.
- Radhamani, K., & Kalaivani, D. (2021). Academic resilience among students: A review of literature. *International Journal of Research and Review*, 8(6), 360-369.
- Sam, R., Jamil, H. B., & Md Zain, A. N. (2024). Academic Adjustment Issues in a Malaysian Research University: The Case of Cambodian, Laotian, Burmese, and Vietnamese Postgraduate Students' Experiences. *Laotian, Burmese, and Vietnamese Postgraduate Students' Experiences (June 01, 2024)*.
- San, C. K., & Guo, H. (2023). Institutional support, social support, and academic performance: Mediating role of academic adaptation. *European Journal of Psychology of Education*, 38(4), 1659-1675.
- Setti, I., Sommovigo, V., & Argentero, P. (2022). Enhancing expatriates' assignments success: the relationships between cultural intelligence, cross-cultural adaptation and performance. *Current Psychology*, 41(7), 4291-4311.

- Sharma, R., & Kumar, P. (2016). Emotional intelligence and stress coping styles: A study of doctors of private hospitals in and around Chandigarh. *International Journal of Management & Social Sciences*, 3(3). <https://doi.org/10.21013/jmss.v3.n3.p24>
- Shuo, Z., Xuyang, D., Xin, Z., Xuebin, C., & Jie, H. (2022). The relationship between postgraduates' emotional intelligence and well-being: the chain mediating effect of social support and psychological resilience. *Frontiers in psychology*, 13, 865025.
- Singh, J. K. N., & Jack, G. (2018). The benefits of overseas study for international postgraduate students in Malaysia. *Higher Education*, 75(4), 607-624.
- Sundaramoorthy, J., Eapen, J. C., Geddam, S., Deepthi, D. P., Pandey, V., Eslavath, R., Yadav, R., & Govindappa, L. (2025). Emotional Intelligence and Cross-Cultural Adaptation of Indian Students in the Context of Interstate Education. *International Journal of Interdisciplinary Cultural Studies*, 20(1), 1-15.
- Vella, S. L. C., & Pai, N. B. (2019). A theoretical review of psychological resilience: Defining resilience and resilience research over the decades. *Archives of Medicine and Health Sciences*, 7(2), 233-239.
- Wagnild, G. M., & Young, H. M. (1993). Development and psychometric. *Journal of nursing measurement*, 1(2), 165-17847.
- Ward, C., & Fischer, R. (2008). Personality, cultural intelligence and cross-cultural adaptation. *Handbook of cultural intelligence: Theory, measurement, and applications*, 159- 173.
- Young, T. J., & Schartner, A. (2014). The effects of cross-cultural communication education on international students' adjustment and adaptation. *Journal of multilingual and multicultural development*, 35(6), 547-562.
- Zawawi, T. S. T. M., Zain, A. D. M., Embong, R., & Alwi, E. A. Z. E. (2022). The Existence Of Religious Tolerance In The Multi-Racial Society of Malaysia. *Journal of Positive School Psychology* <http://journalppw.com>, 6(3), 1976-1982.
- Zhai, X., & Razali, A. B. (2022). International Chinese postgraduate students' adaptation strategies for oral English communication practices in Malaysian higher education institutions. *Education Research International*, 2022(1), 1-11..
- Zhou, Z., & Xiang, J. (2025). The relationship between physical activity and mental toughness among Chinese university students: the chain-mediated role of self-esteem and social support. *Frontiers in Psychology*, 16, 1592192.

Appendix A: List of Instruments Used

Emotional Intelligence (1- Completely strongly disagree, 2- Strongly disagree, 3- Disagree, 4- Neutral, 5- Agree, 6- Strongly agree, 7- Completely strongly agree)

1. Expressing my emotions with words is not a problem for me.
2. I often find it difficult to see things from another person's viewpoint. (-)
3. On the whole, I'm a highly motivated person.
4. I usually find it difficult to regulate my emotions. (-)
5. I generally don't find life enjoyable. (-)
6. I can deal effectively with Malaysian people.
7. I tend to change my mind frequently. (-)
8. Generally, I find it difficult to know exactly what emotion I'm feeling. (-)
9. On the whole, I'm comfortable with the way I look.
10. I often find it difficult to stand up for my rights. (-)
11. I'm usually able to influence the way other people feel.
12. On the whole, I have a gloomy perspective on most things. (-)
13. Those close to me often complain that I don't treat them right. (-)
14. I often find it difficult to adjust my life according to the circumstances. (-)
15. On the whole, I'm able to deal with stress.
16. I often find it difficult to show my affection to those close to me. (-)
17. I'm normally able to "get into someone's shoes" and experience their emotions.
18. I normally find it difficult to keep myself motivated. (-)
19. I'm usually able to find ways to control my emotions when I want to.
20. On the whole, I'm pleased with my life.
21. I would describe myself as a good negotiator.
22. I tend to get involved in things I later wish I could get out of. (-)
23. I'm generally aware of my emotions as I experience them.
24. Given my circumstances, I feel good about myself.
25. I tend to "back down" even if I know I'm right. (-)
26. I don't seem to have any power at all over other people's feelings. (-)
27. I generally believe that things will work out fine in my life.
28. I find it difficult to bond well even with those close to me. (-)
29. Generally, I'm able to adapt to new environments.
30. Others admire me for being relaxed.

Resilience (1- Strongly Disagree, 2- Disagree, 3- Neutral, 4- Agree, 5- Strongly Agree)

1. Although I may not like some of the tasks given to me at university, I would make sure that I do it.
2. I can usually find something to laugh about at university.
3. I can handle many tasks at one time.
4. I can cope with task difficulties by myself.
5. My self-efficacy gets me through hard times in dealing with problems in university.
6. I can do many tasks on my own.
7. I do not have to rely much on my friends at university.
8. I can get through hard times in doing certain tasks by myself.
9. I keep being interested in my course.

10. I take things one day at a time.
11. I have enough energy to do my task efficiently.
12. I try not to overthink about my task load.
13. I try to be creative in solving problems.
14. I make sure I complete important tasks.
15. I usually manage one to do my tasks one way or the other.
16. I feel satisfied when I do my task well.
17. In an emergency at school, I am someone my friends can rely.
18. When I make plans, I follow through them.
19. I am a determined student.
20. My university life has meaning.
21. I do not dwell on things that I cannot do anything about.
22. When I'm in a difficult situation, I still can focus on my task.

Social Support (1- Does not apply, 2- Slightly applicable, 3- Moderately applicable, 4- Highly applicable, 5- Exactly applicable)

1. I can easily find someone in Malaysia who can look after my home when I'm not there.
2. There are people in Malaysia who accept me the way I am without reservations.
3. I receive a lot of understanding and security from others.
4. There is someone very close to me whose help I can always count on.
5. If I need to, I can borrow something from Malaysian friends or neighbours without any problems.
6. I have friends/relatives who will definitively take time to listen if I need someone to talk to.
7. I know several people with whom I like to do things in Malaysia.
8. I have friends/relatives who sometimes simply give me a hug.
9. When I am sick, I can ask friends/relatives to handle important things for me without hesitation.
10. If I'm very depressed, I know who I can turn to.
11. There are people who share both joy and sorrow with me.
12. I have some friends/ relatives with whom I can be quite playful.
13. There is someone close to me in whose presence I feel comfortable without any reservations.
14. There is a group of people (friends, clique) that I belong to and whom I meet often in Malaysia

Cross-cultural Adaptability (1- Strongly Disagree, 2- Disagree, 3- Neutral, 4- Agree, 5- Strongly Agree)

1. I can cope well with adverse circumstances.
2. I feel energetic.
3. I see plenty of interesting challenges.
4. I can cope well with the stress in my life.
5. I have influence over my personal circumstances.
6. I can easily handle setbacks.
7. If something unexpected happens, I can easily adapt.
8. I can cope well with my daily life in Malaysia.

9. If I encounter difficulties, I can find a way out.
10. If I want something, I go for it.

The Mediating Role of Cognitive Reappraisal in The Relationship Between Perceived Autonomy Support and Mental Health Among Vocational College Students in Shanghai, China

ABSTRACT

Authors: Bodier Wang¹, Tajularipin Sulaiman^{1*}, Nuraini Jafri¹, Kai Yan Wong²

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

²Faculty of Education, Languages, Psychology & Music, SEGi University, Malaysia.

*Corresponding author:

tajulas@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

China's Technical and Vocational Education and Training (TVET) system has experienced policy-driven expansion and reform, drawing increasing attention to the mental health of vocational students. Existing research suggests that students in vocational education contexts may experience elevated levels of psychological distress, highlighting the need to examine educational and psychological factors associated with their mental health. Grounded in Self-Determination Theory (SDT), this study examined whether cognitive reappraisal mediates the association between perceived autonomy support and mental health among vocational college students in Shanghai, China. A quantitative descriptive correlational design was adopted, involving 220 students enrolled in vocational colleges. Data were collected using the Autonomy Support Questionnaire (ASQ), the Learning Climate Questionnaire (LCQ), the Emotion Regulation Questionnaire (ERQ), and the Mental Health Continuum–Short Form (MHC-SF). Reliability analysis indicated good to excellent internal consistency across the study measures (Cronbach's $\alpha = .831-.938$). Confirmatory factor analysis supported the adequacy of the measurement model, with acceptable overall fit and evidence of convergent validity. Correlation analysis showed that perceived autonomy support, cognitive reappraisal, and mental health were all positively associated. Mediation analysis further indicated that cognitive reappraisal partially mediated the association between perceived autonomy support and mental health ($\beta = .176$, 95% CI [0.098, 0.259]). The model explained 29.6% of the variance in cognitive reappraisal and 48.3% of the variance in mental health. These findings suggest that autonomy-supportive environments may be linked to better mental health partly through students' use of cognitive reappraisal, with implications for student support and well-being promotion in TVET.

Keywords: Perceived Autonomy Support, Cognitive Reappraisal, Mental Health, Self-Determination Theory, Emotion Regulation, TVET.

1. Introduction

China's vocational education system, situated within the broader Technical and Vocational Education and Training (TVET) sector, has experienced substantial policy support and structural reform over the past decade. Since the release of the *National Medium- and Long-Term Education Reform and Development Outline (2010–2020)* by the State Council, vocational education has been positioned as a strategic component of national human-capital development, with increasing emphasis on the quality and sustainability of vocational talent cultivation.

Within this system, vocational higher education has expanded rapidly. By 2022, China had established a large-scale vocational education infrastructure nationwide, reflecting substantial institutional expansion and student participation (MOE, 2023). However, the rapid expansion of vocational colleges has also drawn attention to student-level outcomes, particularly the psychological well-being of vocational college students, a population that has long remained underrepresented in mainstream educational and psychological research.

Vocational college students must navigate academic coursework, skill-based training, and early career choice during the transition from late adolescence to emerging adulthood, a developmental period characterized by heightened emotional reactivity and increased sensitivity to stress (Ahmed et al., 2015; Orben, 2020). While some studies suggest that vocational college students may also experience concerns related to social comparison and perceived status, empirical evidence in this population remains limited, underscoring the need for more rigorous investigation into the psychosocial factors shaping their mental health.

Against this developmental and contextual backdrop, both teacher and parental support have been identified as important external resources influencing vocational college students' emotional adjustment. Research conducted in vocational college contexts indicates that perceived teacher emotional support is positively associated with mental well-being and related to psychological resilience and anxiety among Chinese vocational college students (Fan & Liu, 2024; Zeng et al., 2019). Similarly, teacher support has been found to enhance students' positive emotions through behavioral and cognitive pathways, such as increased academic control and engagement with peers (Chen & Leung, 2023). Beyond the institutional context, parental autonomy support has been shown to exert a lasting influence on adolescents' and young adults' emotional development. Prior studies demonstrate that parental autonomy support is associated with greater life satisfaction, better mental health, and more adaptive emotion-regulation tendencies (Li & Liu, 2023). Among vocational college students specifically, parental support has been linked to stronger autonomous motivation and healthier adjustment patterns (Shen et al., 2023), while parental emotional and autonomy support has been found to reduce depressive symptoms through psychological satisfaction and cognitive reappraisal (Van der Kaap-Deeder et al., 2020).

However, these relationships remain insufficiently examined in vocational and TVET contexts, particularly among vocational college students in China, whose academic pathways and developmental demands differ in important ways from those of general higher education populations. In the present study, Perceived Autonomy Support (PAS) was examined as an aggregate construct reflecting students' overall perceptions of autonomy support across important interpersonal contexts. This approach was adopted for both practical and

psychometric reasons, as it allowed the analysis to focus on the broader autonomy-supportive climate experienced by students while remaining consistent with the study's overall measurement and analytic framework.

Cognitive Reappraisal (CR), an emotional regulation strategy involving the reframing of stressful events to modify emotional responses. Evidence consistently shows that CR is associated with reduced negative affect, and improved mental health outcomes among adolescents and emerging adults (Gross, 2015; Shum et al., 2025; Roos & Bennett, 2023). Recent studies with Chinese college populations further indicate that frequent use of CR predicts lower stress responses and higher well-being (Fan & Liu, 2024). Given the specific psychological vulnerabilities of vocational college students, CR may serve as a crucial internal resource that enables them to translate PAS into healthier emotional and mental states.

CR may be particularly relevant in East Asian contexts, where emotional moderation, self-regulation, and the maintenance of interpersonal harmony are often culturally valued (Song et al., 2024; Wei et al., 2013). In such contexts, reinterpreting stressful situations in less disruptive ways may be more socially reinforced than more outwardly expressive responses. This cultural consideration is relevant to the present study, as it suggests that the mediating role of CR among Chinese vocational college students should be interpreted not only as an individual regulatory process, but also as one shaped by broader sociocultural expectations.

Grounded in Self-Determination Theory (SDT) and emotion-regulation theory, this study pre-specified a mediation model in which PAS was hypothesised to predict Mental Health (MH) both directly and indirectly through cognitive reappraisal among vocational college students in Shanghai, China. Specifically, PAS was expected to positively predict CR and MH, while CR was expected to positively predict MH. The mediation test was treated as the primary confirmatory analysis, whereas supplementary analyses assessed robustness after adjusting for age and gender. The findings aim to inform educational practice and MH support in vocational college contexts, particularly within Shanghai and similar TVET settings in China.

One contribution of the present study is its focus on vocational college students, who remain underrepresented in mental-health research within China's TVET context. It also examines whether CR mediates the association between PAS and MH. The results clarify how CR is involved in the relationship between PAS and MH and offer implications for student support and mental health promotion in vocational higher education.

Research Questions:

RQ1: What are the relationships between perceived autonomy support, cognitive reappraisal, and mental health among vocational college students in Shanghai?

RQ2: Does Cognitive Reappraisal mediate between perceived autonomy support and mental health among vocational college students?

The hypothesised relationships among the study variables are reflected in the following research hypotheses:

Research Hypotheses:

Ha₁: Perceived autonomy support has a positive effect on cognitive reappraisal.

Ha₂: Cognitive reappraisal has a positive effect on mental health.

Ha₃: Perceived autonomy support has a positive effect on mental health.

Ha₄: Cognitive reappraisal mediates the relationship between perceived autonomy support and mental health.

2. Literature Review

2.1. Emotion Regulation

Cognitive Reappraisal (CR), an emotion regulation strategy that involves changing the way one interprets emotion-eliciting situations to alter their emotional impact (Gross, 1998). Unlike suppression, which inhibits emotional expression, CR targets cognitive evaluation, allowing individuals to reduce negative emotions while enhancing positive emotions (Gross, 2015). CR has been associated with better psychological adjustment, improved interpersonal relationships, and higher subjective well-being (Fan & Liu, 2024).

A growing body of research suggests that CR not only attenuates negative emotional experiences but also supports more effective coping with stress (Ahmed et al., 2015; Fan & Liu, 2024). Empirical research with adolescents and young adults indicates that reappraising stressful academic and developmental challenges in more constructive terms is associated with better adjustment under stress (Troy et al., 2010; Silvers, 2022). Mechanistically, such reframing can reduce threat appraisals and help individuals maintain approach-oriented engagement with valued goals rather than disengaging or ruminating. In TVET contexts, where students often face sustained performance pressure and career uncertainty, CR may therefore be particularly beneficial. Consistent with this account, more frequent use of CR has been linked to lower anxiety, depression, and burnout, as well as higher life satisfaction and resilience (Riepenhausen et al., 2022; Stover et al., 2024; Chen & Leung, 2023).

CR is also shaped by cultural and educational contexts. Prior evidence suggests that Chinese individuals report relatively frequent use of cognitive reappraisal, and East Asian educational settings are often characterised by strong academic and social evaluative pressures, which may increase the relevance of reappraisal in managing school related stress (Qu et al., 2017; Chyu et al., 2022). A plausible explanation is that reappraisal aligns with culturally valued self-regulation goals, such as maintaining composure and interpersonal harmony, and can be enacted without overt emotional expression. For vocational college students facing academic and career challenges, CR provides a critical cognitive resource to enhance adaptation and psychological well-being (Li & Liu, 2023).

Neurocognitive studies further support the effectiveness of CR. Functional imaging and psychophysiological research indicate that CR engages prefrontal regulatory circuits and modulates amygdala activity, reducing emotional reactivity and enhancing emotion regulation capacity (Ahmed et al., 2015; Morawetz et al., 2017). This evidence demonstrates that CR is not only a psychological strategy but also has measurable neurocognitive correlates, reinforcing its role as a robust factor in improving Mental Health (MH). These neurocognitive findings are consistent with the mechanistic view that reappraisal involves top-down regulatory processes that can dampen emotional reactivity.

Recent literature indicates that CR may function as an important mediating pathway between environmental support and psychological outcomes. Within the framework of Self-Determination Theory (SDT), autonomy-supportive environments are expected to satisfy students' basic psychological needs, especially the need for autonomy, thereby fostering more autonomous rather than controlled forms of regulation (Deci & Ryan, 1985, 2000; Ryan & Deci, 2024). When students experience their environment as autonomy-supportive, they may be more likely to process stressful situations in flexible, self-endorsed, and reflective ways, rather than in rigid or externally pressured ways. From the perspective of emotion-regulation theory, CR involves reinterpreting potentially stressful situations in ways that modify their emotional impact (Gross, 1998; Gross, 2015). Accordingly, CR can be understood as a theoretically plausible mechanism through which perceived autonomy support may be associated with better mental health. Simultaneously, the present study conceptualises CR as one theoretically grounded mediating pathway, rather than the only mechanism linking autonomy support to MH, as other SDT-relevant processes, such as competence, relatedness, and broader psychosocial support, may also be involved (Ryan & Deci, 2024; Vasconcellos et al., 2020).

2.2. The Relationship between Cognitive Reappraisal, Perceived Autonomy Support and Mental Health

Besides SDT, the environmental and personal resource perspective also provides a useful framework for understanding the connection between situational support, individual self-regulation, and well-being (Gee & Cohodes, 2023; Fan & Liu, 2024). Within this framework, PAS can be considered a key situational resource, while CR is an important individual self-regulation resource associated with better MH outcomes. SDT typically understands autonomy support as an atmosphere or style of interpersonal interaction (Ryan & Deci, 2000). However, in students' daily lives, autonomy support may come from multiple significant others, including teachers, parents, friends and school peers, and these sources may function in distinct ways (Tilga et al., 2021; Audet et al., 2021). Mechanistically, supportive interactions may foster more autonomous and integrative forms of regulation, thereby increasing the likelihood of adaptive meaning construction strategies such as CR (Ryan & Deci, 2000; Benita, 2020).

However, the significance of autonomy support may vary across stressful contexts. Teacher autonomy support is typically embedded in learning environments and is closely tied to task engagement, instructional feedback, and students' perceived control in academic settings, whereas parental autonomy support may be especially salient in relation to family expectations and developmental decision-making, particularly in collectivist cultural contexts in which family-related goals remain highly influential (Marbell-Pierre et al., 2019; Mammadov et al., 2023). By contrast, peers may play a more prominent role in interpersonal coping, as peer interactions can facilitate emotion regulation by providing social reappraisal, validation, and alternative perspectives on emotionally salient situations (Sahi et al., 2023). Although some studies have examined autonomy support from multiple sources, including teachers, parents, and peers, the source-specific mechanisms through which these forms of support operate remain insufficiently differentiated. Accordingly, although the present study operationalizes autonomy support at the overall perceived level, distinguishing among different sources and testing their unique pathways represents an important direction for future research.

Perceived Autonomy Support (PAS) refers to students' perceptions that significant figures in their environment, such as teachers, parents, and peers, encourage choice, self-initiation, and independent problem-solving (Ryan & Deci, 2000). As a core concept of SDT, PAS supports the fulfillment of basic psychological needs, autonomy, competence, and relatedness, which are essential for optimal psychological functioning (Deci & Ryan, 1985; Ryan & Deci, 2024).

Research has consistently shown that autonomy-supportive teaching and parenting are associated with better MH outcomes, including greater well-being and fewer emotional problems, among adolescents and young adults (Ma et al., 2022; Kristensen & Jenø, 2024). In educational settings characterised by substantial academic and developmental demands, PAS may help buffer stress and strengthen students' emotional and motivational resources (Lan & Moscardino, 2019).

Beyond direct effects, PAS also promotes adaptive emotion regulation strategies, particularly CR. Students in autonomy-supportive environments are more likely to engage in cognitive reappraisal, reinterpreting challenging situations in ways that are associated with lower emotional distress (Shen et al., 2023). Previous studies using Structural Equation Modeling (SEM) have reported significant associations among autonomy-supportive environments, CR, and MH, suggesting that CR may play a linking role between PAS and emotional well-being (Li & Liu, 2023; Zeng et al., 2019).

Additionally, PAS may not operate in isolation, but rather in conjunction with other psychosocial resources, such as perceived competence and peer support, in facilitating students' use of adaptive emotion regulation strategies, particularly CR. From the perspective of SDT, autonomy-supportive environments promote more self-endorsed functioning and are closely associated with the satisfaction of basic psychological needs, especially competence, thereby providing a favorable psychological basis for flexible and constructive regulation of stressful experiences (Ryan & Deci, 2000). Empirical research has further shown that perceived teacher autonomy support is positively associated with students' self-efficacy, and that peer support may strengthen the beneficial effects of autonomy-supportive contexts on student functioning (Zhao & Qin, 2021). Related evidence further suggests that supportive interpersonal and psychosocial resources are associated with greater use of CR and may contribute to better stress adaptation and resilience among young people (Mohan et al., 2024; Ye et al., 2022).

In parallel, a growing body of evidence indicates that cognitive reappraisal is a protective regulatory strategy that is positively related to stress adaptation and resilience (Riepenhausen et al., 2022; Stover et al., 2024). In summary, these findings underscore the importance of considering both supportive environments and individual cognitive strategies when examining MH outcomes. While the present study focuses on CR as a theoretically grounded mediator, other mechanisms may also link autonomy support to MH. For example, perceived competence, relatedness need satisfaction, and broader social support may operate as parallel or complementary mediators. Future research could therefore test multi-mediator models to clarify the unique and shared contributions of these pathways.

Within the SDT framework, autonomy-supportive environments are expected to foster more autonomous rather than controlled forms of regulation (Ryan & Deci, 2000; Ryan & Deci, 2024). From this perspective, CR may be understood as a relatively adaptive and self-endorsed emotion-regulation strategy, because it involves actively reinterpreting stressful experiences in

ways that preserve goal-directed functioning rather than merely suppressing emotional expression or reacting under external pressure (Gross, 1998; Gross, 2015; Benita, 2020). This makes CR a theoretically plausible mediator between PAS and MH. Specifically, when students perceive their social environment as autonomy-supportive, they may be more likely to adopt flexible and self-endorsed regulatory responses, which in turn may contribute to better MH outcomes (Fan & Liu, 2024; Shen et al., 2023; Li & Liu, 2023). Simultaneously, the present study does not assume that CR is the only pathway through which PAS may relate to MH. In line with SDT, other mechanisms, such as competence-related beliefs, relatedness need satisfaction, and broader supportive resources, may also operate alongside CR (Ryan & Deci, 2024; Zhao & Qin, 2021; Fan & Liu, 2024).

3. Methodology

3.1. Research Design

This study employed a quantitative descriptive correlational survey design to examine the relationships among perceived autonomy support, cognitive reappraisal, and mental health among vocational higher education students in Shanghai, China. A correlational approach is appropriate for examining associations among psychological constructs that cannot be manipulated experimentally (Creswell & Creswell, 2017). The study was guided by a post-positivist paradigm, which assumes that reality can be approximated through systematic inquiry while acknowledging that complete objectivity is unattainable (Phillips & Burbules, 2000). In line with the study's theoretical framework, mediation analysis was used to examine whether cognitive reappraisal statistically mediated the association between perceived autonomy support and mental health. However, because the data were cross-sectional, the mediation findings should be interpreted as associational rather than as evidence of strict causal relationships.

3.2. Research Sample

Convenience sampling was used to recruit participants from vocational colleges in Shanghai, China, between 2 August 2025 and 26 November 2025. To be eligible for inclusion, participants had to be vocational college students in Shanghai, be at least 18 years old, and participate voluntarily in the study. Only fully completed questionnaires from eligible participants were included in the final analysis. Responses were excluded if the respondent was not a vocational college student in Shanghai, was under 18 years of age, submitted an incomplete questionnaire, or submitted duplicate responses.

The research team first contacted school administrators by email to seek permission and facilitate access. Subsequently, homeroom teachers distributed an online survey link hosted on Questionnaire Star (Wenjuanxing) to eligible students, who completed the questionnaire voluntarily. An a priori power analysis was conducted using G*Power based on a linear multiple regression model, assuming a medium effect size ($f^2 = .15$), an alpha level of .05, a statistical power of .80, and two predictors. The analysis indicated a minimum required sample size of 68 participants. This analysis was used as a general adequacy check for the planned regression-based analyses. In addition, a post hoc sensitivity analysis was conducted in G*Power based on the final sample size of 220, $\alpha = .05$, power = .80, and two predictors. The

analysis indicated that the present sample was sufficient to detect a minimum effect size of $f^2 = .044$, providing further support for the adequacy of the sample for the planned analyses.

A total of 260 questionnaires were distributed during the recruitment period, and 220 completed questionnaires were retained in the final dataset, yielding a response rate of 84.62%. The final sample consisted of 220 vocational college students. All quantitative data were analysed using IBM SPSS Statistics 27.0. Table 1 presents the demographic profile of the participants. Because participants were recruited through convenience sampling from vocational colleges in Shanghai, the findings should be interpreted with caution in relation to generalisability beyond this specific context.

Table 1. Demographic Profile of Students (n=220)

Demographic Variable	Profile	Frequency	Percentage (%)
Gender	Male	132	60.00
	Female	88	40.00
Grade	Grade 1 / First-year	91	41.36
	Grade 2 / Second-year	59	26.82
	Grade 3 / Third-year	70	31.82
Ethnicity	Han	188	85.45
	Others	32	14.55
Total		220	100.00

3.3. Ethical Consideration

This study received ethical approval from the Ethics Committee of Universiti Putra Malaysia (JKEUPM-2025-359).

3.4. Instrument

All instruments used in this study, including the ASQ, LCQ, ERQ, and MHC-SF, were administered in accordance with ethical guidelines and standard research procedures. The instruments were translated from English into Chinese by two bilingual experts with backgrounds in educational psychology. To improve linguistic accuracy and contextual appropriateness, four additional experts, including an educational psychologist, a vocational education specialist, a Chinese language teacher, and an English language teacher, reviewed the translated items and recommended revisions where necessary. After these revisions had been incorporated, the final Chinese version was further reviewed to confirm its suitability for the present study. A formal back-translation procedure was not conducted in the present study, which should be acknowledged as a limitation of the instrument adaptation process.

3.4.1. Emotional Regulation Questionnaire (ERQ)

The ERQ assesses habitual use of cognitive reappraisal and expressive suppression. In the present study, only the six items measuring cognitive reappraisal were used. Responses were rated on a 7-point Likert scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”), with higher scores indicating greater use of cognitive reappraisal. The cognitive reappraisal subscale of the ERQ has demonstrated good reliability in previous research ($\alpha = .79$; Gross & John, 2003).

Perceived autonomy support (PAS) was assessed using two validated instruments: The Autonomy Support Questionnaire (ASQ) and the Learning Climate Questionnaire (LCQ).

3.4.2. Autonomy Support Questionnaire (ASQ)

The ASQ, originally developed by Deci et al. (2006) and later adapted by Legate, Ryan, and Weinstein (2012), is a self-report measure assessing perceived autonomy support from parents, friends, and school peers. It comprises three subscales: parents (ASQP), friends (ASQF), and school peers (ASQSP), with a total of 30 items. Responses were rated on a 7-point Likert scale ranging from 1 (“not at all true”) to 7 (“very true”). The scale has demonstrated good reliability in previous research ($\alpha = .89$; Legate et al., 2012). Although the ASQ was originally developed in a Western context, the translated version used in the present study underwent expert review to improve its linguistic accuracy and contextual appropriateness. Moreover, the scale demonstrated acceptable psychometric properties in the current sample. However, it remains possible that certain culturally specific patterns of interpersonal support and relational expectations in the Chinese TVET context were not fully represented.

3.4.3 Learning Climate Questionnaire (LCQ)

The LCQ was used to assess students’ perceptions of teacher autonomy support. The full 15-item version was adopted, and responses were rated on a 7-point Likert scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”), including one reverse-scored item. The LCQ has demonstrated good reliability in previous research ($\alpha = .82$; Williams & Deci, 1996). Although both the ASQ and LCQ used 7-point response formats, their scale anchors differed, which should be taken into account when interpreting descriptive comparisons across the two measures.

3.4.4 Mental Health Continuum – Short Form (MHC-SF)

The MHC-SF is a 14-item questionnaire that assesses mental health across three domains: emotional, psychological, and social well-being. It was developed as a brief version of the original Mental Health Continuum long form (Keyes, 2009). Participants indicated the frequency of each experience over the past month on a 6-point Likert scale ranging from 1 (“never”) to 6 (“every day”). The MHC-SF has demonstrated good reliability in previous research ($\alpha = .80$; Keyes, 2009).

3.5. Data Analysis

Data analysis was conducted using IBM SPSS Statistics 27.0 and AMOS 28.0. Descriptive statistics, reliability analysis, and Pearson correlation analysis were performed in SPSS 27.0. Confirmatory Factor Analysis (CFA) was conducted in AMOS 28.0 to examine the measurement properties of the study constructs.

Scale scores for each construct were calculated using raw mean scores across the relevant items rather than standardized z-scores. Because the instruments used different response formats, score interpretation was based on each scale’s original response range, and direct comparison of mean values across different instruments was made with caution. Cronbach’s alpha

coefficients were computed to assess internal consistency. CFA was used to evaluate the measurement model, and model fit was assessed using multiple indices, including χ^2/df , RMSEA, IFI, CFI, TLI, NFI, and RFI. In addition, composite reliability (CR) and average variance extracted (AVE) were calculated to assess convergent validity.

After the measurement model had been evaluated, descriptive statistics and Pearson correlation coefficients were computed for the main study variables. Mediation analysis was then conducted in AMOS 28.0 using a structural path model to examine whether cognitive reappraisal mediated the association between perceived autonomy support and mental health. In addition to the primary mediation model, an adjusted model was estimated with age and gender included as control variables to examine the robustness of the observed associations. Prior to interpreting the adjusted model, collinearity diagnostics were conducted using tolerance and variance inflation factor (VIF) values to assess whether multicollinearity among the predictors was a concern. The indirect effect was estimated using a bias-corrected bootstrap procedure with 5,000 resamples and 95% confidence intervals. An indirect effect was considered statistically significant when the confidence interval did not include zero. Because the data were cross-sectional, the mediation results were interpreted as statistical associations rather than as evidence of strict causal relationships.

4. Findings

4.1. Measurement Model Assessment

Cronbach's alpha (α) coefficients were calculated to assess the internal consistency of all study measures. As shown in Table 2, the study measures demonstrated acceptable to excellent reliability and convergent validity. The overall mental health construct showed good internal consistency ($\alpha = .847$), the cognitive reappraisal subscale of the ERQ showed high reliability ($\alpha = .892$), and the overall perceived autonomy support construct demonstrated excellent internal consistency ($\alpha = .925$). In addition, perceived teacher autonomy support, as measured by the LCQ, showed excellent internal consistency ($\alpha = .938$). These results indicate that the measures were sufficiently reliable for the subsequent analyses. Notably, the Cronbach's alpha for the ERQ cognitive reappraisal subscale in this study (.892) was higher than that reported in the original scale development study (.79; Gross & John, 2003), which may reflect differences related to the adapted Chinese version or the characteristics of the present sample.

Confirmatory Factor Analysis (CFA) was conducted in AMOS 28.0 to examine whether the established factor structures of the study measures were replicated in the present sample. The measurement model showed satisfactory overall fit to the data, with $\chi^2/\text{df} = 1.812$, RMSEA = .048, IFI = .967, CFI = .966, TLI = .950, NFI = .929, and RFI = .896. Most fit indices met commonly used criteria for acceptable model fit. Although the RFI value (.896) was slightly below the conventional .90 thresholds, it was very close to the range typically interpreted as acceptable for incremental fit indices, and the remaining fit indices consistently supported acceptable to good overall model fit (Hu & Bentler, 1999; Schermelleh-Engel et al., 2003). As shown in Table 2, all reported standardized loadings were statistically significant and exceeded .50, indicating adequate indicator reliability (Hair et al., 2019). Composite reliability (CR) values ranged from .703 to .861, exceeding the commonly recommended threshold of .70 (Hair et al., 2019), and average variance extracted (AVE) values ranged from .563 to .624,

exceeding the recommended criterion of .50, thereby supporting convergent validity (Fornell & Larcker, 1981). Discriminant validity was further assessed using the Fornell–Larcker criterion (Fornell & Larcker, 1981). As shown in Table 3, the square root of the average variance extracted for each construct exceeded its correlations with the other constructs. Specifically, the square root of the AVE was 0.760 for mental health, 0.790 for cognitive reappraisal, and 0.750 for perceived autonomy support, all of which were greater than the corresponding inter-construct correlations. These results support the discriminant validity of the measurement model. In summary, these results indicate that the measurement model was appropriate for the subsequent analyses.

Table 2. Measurement Properties of Study Constructs

Construct		Cronbach's α	Factor Loading	Composite Reliability (CR > .70)	Average Variance Extracted (AVE > .50)
Mental Health (MH)		0.847		0.703	0.578
EWB			0.63		
SWB			0.77		
PWB			0.75		
Cognitive Reappraisal (CR)		.892		0.861	0.624
ERQ1			0.76		
ERQ2			0.82		
ERQ3			0.81		
ERQ4			0.85		
ERQ5			0.71		
ERQ6			0.78		
Perceived Autonomy Support (PAS)		0.925		0.744	0.563
Parents			0.79		
Friends			0.73		
School Peers			0.76		
Teachers			0.72		

Table 3. Discriminant Validity Assessment

Construct	1	2	3
Mental Health (MH)	0.760		
Cognitive Reappraisal (CR)	0.582	0.790	
Perceived Autonomy Support (PAS)	0.645	0.481	0.750

Note. Diagonal values in bold represent the square roots of the AVE for each construct. Off-diagonal values represent inter-construct correlations.

4.2. Descriptive Statistics and Correlation Analysis

Descriptive statistics and bivariate correlations for the main study variables are presented in Table 4. On the 1–6 MHC-SF scale, the mean mental health score was 4.43 ($SD = 0.92$). On the 1–7 scales, the mean score was 5.65 ($SD = 1.29$) for cognitive reappraisal and 5.31 ($SD = 0.96$) for perceived autonomy support. Because the instruments used different response formats, the reported means should be interpreted with reference to each scale's original response range, and direct comparisons of mean values across different instruments should be made with

caution. Pearson correlation analysis indicated that the three main study variables were all significantly and positively correlated at the .01 level. Mental health was positively associated with cognitive reappraisal ($r = .582, p < .01$) and perceived autonomy support ($r = .645, p < .01$). Cognitive reappraisal was also positively associated with perceived autonomy support ($r = .481, p < .01$). Overall, these findings indicate that higher levels of perceived autonomy support and cognitive reappraisal were associated with better reported mental health, and that perceived autonomy support was also positively associated with cognitive reappraisal. These results provided the basis for the subsequent mediation analysis.

Table 4. Descriptive statistics and correlation analysis (n=220)

Construct	Mean	SD	1	2	3
Mental Health	4.43	0.92	-		
Cognitive Reappraisal	5.65	1.29	0.582**	-	
Perceived Autonomy Support	5.31	0.96	0.645**	0.481**	-

Note. $p < .01$ (two-tailed). Mental Health was measured on a 1–6 scale; Cognitive Reappraisal and Perceived Autonomy Support were measured on 1–7 scales. Means should therefore be interpreted with reference to each scale's response range.

4.3. Mediation Analysis

Figure 1 and Table 5 present the results of the AMOS structural path model testing cognitive reappraisal as a mediator of the association between perceived autonomy support and mental health. Perceived autonomy support was positively associated with cognitive reappraisal ($\beta = 0.544, SE = 0.057, p < .001$), and cognitive reappraisal was positively associated with mental health ($\beta = 0.323, SE = 0.058, p < .001$). The total effect of perceived autonomy support on mental health was $\beta = 0.640$. When cognitive reappraisal was included in the model, the direct effect of perceived autonomy support on mental health remained significant ($\beta = 0.464, SE = 0.058, p < .001$), indicating partial mediation. Bootstrap mediation analysis based on 5,000 resamples indicated a significant indirect effect of perceived autonomy support on mental health through cognitive reappraisal ($\beta = 0.176, 95\% \text{ CI } [0.098, 0.259]$). Because the confidence interval did not include zero, the indirect effect through cognitive reappraisal was considered statistically significant. The model explained 29.6% of the variance in cognitive reappraisal ($R^2 = .296$) and 48.3% of the variance in mental health ($R^2 = .483$).

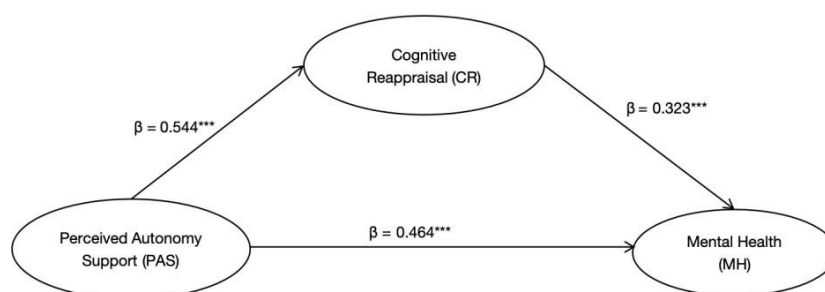


Figure 1. Mediation Effects Model

Table 5. Mediation Effects of CR Between PAS and MH

Path	Beta (β)	SE	t	p	95% CI (Bootstrap)
PAS $\rightarrow$ CR	0.544	0.057	9.541	<.001	-
CR $\rightarrow$ MH	0.323	0.058	5.543	<.001	-
PAS $\rightarrow$ MH (direct effect)	0.464	0.058	7.968	<.001	-
PAS $\rightarrow$ CR $\rightarrow$ MH (indirect effect)	0.176	-	-	-	[0.098, 0.259]
PAS $\rightarrow$ MH (total effect)	0.640	-	-	-	-
R ² for CR	0.296				
R ² for MH	0.483				

Note. PAS = Perceived Autonomy Support; CR = Cognitive Reappraisal; MH = Mental Health. Bootstrap samples = 5,000.

4.4. Common-Method Bias Check

Given the cross-sectional and self-report nature of the data, an additional assessment of common-method bias was conducted using the full-collinearity VIF approach. Prior methodological work has suggested that full-collinearity VIF values below 3.3 indicate that common-method bias is unlikely to be severe (Kock, 2015). In the present study, the full-collinearity VIF values were 1.836 for mental health, 1.627 for cognitive reappraisal, and 1.716 for perceived autonomy support. These results therefore suggest that common-method bias was unlikely to pose a serious threat to the present findings.

Table 6. Full-Collinearity VIF Values for Common-Method Bias Assessment

Construct	VIF
Mental Health	1.836
Cognitive Reappraisal	1.627
Perceived Autonomy Support	1.716

Note. All full-collinearity VIF values were below 3.3, indicating that common-method bias was unlikely to be a serious concern (Kock, 2015).

To examine the robustness of the mediation findings after accounting for basic demographic characteristics, an adjusted model including age and gender was also considered. In addition, conventional collinearity diagnostics were conducted for the adjusted regression model, with mental health as the dependent variable and perceived autonomy support, cognitive reappraisal, age, and gender as predictors. As shown in Table 7, the tolerance values ranged from .738 to .980, and the Variance Inflation Factor (VIF) values ranged from 1.021 to 1.354. These values were well within commonly accepted limits, according to which tolerance values below .20 and VIF values above 5 may indicate problematic multicollinearity (Hair et al., 2010; O'Brien, 2007). Therefore, multicollinearity was not considered a serious concern in the adjusted model.

Table 7. Multicollinearity Statistics for the Adjusted Model

Variables	Tolerance	Collinearity VIF
Perceived Autonomy Support (PAS)	0.744	1.344
Cognitive Reappraisal (CR)	0.738	1.354
Gender	0.980	1.021
Age	0.951	1.051

Note. Dependent variable = mental health (MH). Predictors = PAS, CR, Age, and Gender.

4.5 Hypothesis Testing Results

Table 8 summarizes the results of hypothesis testing. Ha₁, Ha₂, and Ha₃ were supported, as perceived autonomy support significantly and positively predicted cognitive reappraisal, cognitive reappraisal significantly and positively predicted mental health, and perceived autonomy support significantly and positively predicted mental health. Ha₄ was also supported, as cognitive reappraisal significantly mediated the relationship between perceived autonomy support and mental health. Overall, the findings supported the proposed mediation model.

Table 8. Summary of Hypothesis Testing

Research Hypotheses	Results
Ha ₁ : Perceived autonomy support has a positive effect on cognitive reappraisal.	Supported
Ha ₂ : Cognitive reappraisal has a positive effect on mental health.	Supported
Ha ₃ : Perceived autonomy support has a positive effect on mental health.	Supported
Ha ₄ : Cognitive reappraisal mediates the relationship between perceived autonomy support and mental health.	Supported

5. Discussion

The present study investigated the relationships among PAS, CR, and MH among vocational college students in Shanghai, and further examined whether CR mediated the relationship between perceived autonomy support and mental health. Guided by SDT (Deci & Ryan, 2000) and the process model of emotion regulation (Gross, 1998; 2015), the findings provide new insights into how perceived autonomy support and students' use of cognitive reappraisal interactively influence mental health among vocational college students.

5.1. Relationships Among Perceived Autonomy Support, Cognitive Reappraisal, and Mental Health

RQ1: What are the relationships between Perceived Autonomy Support, Cognitive Reappraisal, and Mental Health among vocational college students in Shanghai?

For RQ1, perceived autonomy support, reflecting support from parents, friends, school peers, and teachers, was positively associated with mental health among vocational college students in Shanghai, China. Specifically, perceived autonomy support showed a significant positive correlation with mental health ($r = .645, p < .01$), suggesting that students who reported higher levels of autonomy support also tended to report better emotional, psychological, and social well-being. This finding is consistent with self-determination theory, which posits that autonomy-supportive environments facilitate the satisfaction of basic psychological needs for autonomy, competence, and relatedness, thereby promoting psychological functioning and well-being (Ryan & Deci, 2024). Empirical studies further indicate that college students who perceive higher autonomy support tend to report greater intrinsic motivation and life satisfaction, and related evidence has also linked autonomy-supportive contexts to positive motivational and adaptive outcomes in higher education (Black & Deci, 2000; Lozano-Jiménez et al., 2021; Vasconcellos et al., 2020).

However, the relatively strong association between perceived autonomy support and mental health should also be interpreted cautiously. In addition to reflecting a substantively meaningful relationship, the correlation may partly capture shared variance arising from the use of self-report measures, common method bias, or some conceptual proximity between supportive interpersonal experiences and positive psychological functioning. Even so, the measurement results provide preliminary support for treating these constructs as empirically distinguishable in the present study.

The findings also showed that cognitive reappraisal was positively associated with both perceived autonomy support and mental health. Cognitive reappraisal was significantly correlated with mental health ($r = .582, p < .01$), while perceived autonomy support was significantly correlated with cognitive reappraisal ($r = .481, p < .01$). These results are consistent with Gross's (2015) emotion regulation framework, which identifies cognitive reappraisal as an adaptive strategy associated with reduced distress and greater life satisfaction. Recent evidence similarly indicates that reappraisal may protect adolescents from stress, anxiety, and depressive symptoms (Gee & Cohodes, 2023). Taken together, the findings suggest that autonomy-supportive environments may be linked to better mental health both directly and through their positive association with adaptive emotion-regulation processes (Ryan & Deci, 2024).

5.2. Mediating Role of Cognitive Reappraisal

RQ2: Does CR mediate between PAS and MH among vocational college students?

For RQ2, mediation analysis revealed that cognitive reappraisal significantly mediated the association between perceived autonomy support and mental health. Specifically, perceived autonomy support positively predicted cognitive reappraisal ($\beta = 0.544, p < .001$), cognitive reappraisal positively predicted mental health ($\beta = 0.323, p < .001$), and the indirect effect of perceived autonomy support on mental health through cognitive reappraisal was significant ($\beta = 0.176, 95\% \text{ CI } [0.098, 0.259]$). Because the bootstrap confidence interval did not include zero, the mediating effect of cognitive reappraisal was considered statistically significant. At the same time, the direct effect of perceived autonomy support on mental health remained significant after cognitive reappraisal was included in the model ($\beta = 0.464, p < .001$), indicating partial mediation rather than full mediation.

These findings are consistent with theoretical expectations. Autonomy support is thought to promote flexible thinking, volitional engagement, and adaptive coping, all of which may facilitate cognitive reappraisal (Ryan & Deci, 2024). Adolescents who perceive support from significant others may feel psychologically safer and more able to reinterpret stressors in constructive ways, thereby enhancing emotional competence and reducing vulnerability to distress. Empirical studies likewise suggest that autonomy-supportive environments are positively associated with social emotional competence and adaptive emotion regulation related skills, including cognitive reappraisal (Collie et al., 2024). Consistent with these perspectives, the present findings suggest that perceived autonomy support may be associated with better mental health partly through more adaptive emotion-regulation processes among vocational college students.

Importantly, the finding of partial mediation indicates that cognitive reappraisal represents one meaningful pathway, but not the only pathway, through which perceived autonomy support may be associated with mental health. The remaining direct effect suggests that other mechanisms may also be involved. Within the self-determination theory framework, these mechanisms may include basic psychological need satisfaction, perceived competence, relatedness, or broader forms of social support. Thus, autonomy-supportive environments may be linked to mental health both indirectly, through adaptive emotion regulation, and directly, through broader motivational and relational processes.

Overall, the findings extend previous literature by showing that cognitive reappraisal plays a significant, but partial, mediating role in the association between perceived autonomy support and mental health in a vocational college context. This is particularly relevant in TVET settings, where students often face concurrent academic, interpersonal, and career-related pressures.

5.3. Implications

The present findings carry several practical implications. First, they underscore the importance of fostering autonomy-supportive environments across both family and school contexts. Recent work has shown that practices such as taking students' perspectives, providing explanatory rationales, using less controlling language, and supporting meaningful choice are characteristic of autonomy-supportive instruction and are associated with stronger motivation, need satisfaction, and well-being-related outcomes (Kong et al., 2024; Zachariou et al., 2024). Accordingly, training initiatives for teachers and parents may benefit from explicitly incorporating autonomy-supportive interaction strategies. This implication is further supported by intervention and review evidence showing that autonomy-supportive practices can be intentionally cultivated and can produce beneficial effects on students' motivational and adaptive outcomes (Su & Reeve, 2011; Reeve & Cheon, 2021).

Second, the significant mediating role of cognitive reappraisal suggests that interventions aimed at strengthening emotion-regulation capacities may be especially relevant for vocational college students. Meta-analytic evidence indicates that reappraisal-based training, mindfulness-based programs, and cognitive-behavioral approaches can improve psychological well-being and resilience among adolescents and young adults (Gee & Cohodes, 2023; Dunning et al., 2019). Embedding structured emotional skills training within vocational education may therefore help students cope more effectively with academic and transitional stressors. In practice, this may include integrating modules on stress appraisal, emotional awareness, and adaptive cognitive reframing into student support services and mental health promotion programs.

Third, the findings suggest that vocational colleges should consider strengthening autonomy-supportive pedagogy and student-support systems at the institutional level. Educational practices that encourage student voice, support self-initiation, and communicate in non-controlling ways may help promote both adaptive emotion regulation and better mental health. Professional development focused on relational pedagogy and supportive communication may therefore be particularly valuable in TVET settings. This implication is also consistent with recent educational scholarship indicating that autonomy-supportive teaching is malleable,

beneficial, and capable of improving students' engagement, motivation, and broader adaptive functioning (Reeve & Cheon, 2021).

Finally, the present findings are consistent with recent national policy trends in China that emphasize strengthening school mental health systems, enhancing teacher training, and integrating emotion regulation education into the curriculum (MOE, 2023). In this respect, the study provides empirical support for coordinated approaches that combine interpersonal support with psychological skill-building in vocational and higher education contexts. Recent review evidence from China further indicates that national attention to school mental health has been accompanied by a growing body of prevention and intervention efforts across educational settings, including colleges and universities (Qu et al., 2024).

5.4. Limitations and Future Research

The reliance on self-report measures may introduce response biases and shared method variance, which may have inflated the observed associations among perceived autonomy support, cognitive reappraisal, and mental health. Future studies could integrate behavioral, peer-reported, teacher-reported, or physiological indicators of emotion regulation to enhance validity and reduce common method bias (Podsakoff et al., 2024).

In addition, the present findings should be interpreted with caution because the data were cross-sectional. Although the hypothesized mediation model was theoretically grounded, cross-sectional mediation analysis cannot establish definitive causal or temporal ordering among the variables. Therefore, the mediating role of cognitive reappraisal should be understood as a statistical association rather than conclusive evidence of causal mediation. Future studies could adopt longitudinal, cross-lagged, or experimental designs to provide stronger evidence regarding the temporal relationships among perceived autonomy support, cognitive reappraisal, and mental health (Maxwell & Cole, 2007; Maier et al., 2023).

The sample was drawn from a single city, which may limit generalizability to other vocational or higher education contexts. In addition, participants were recruited through convenience sampling and teacher-mediated distribution of the survey link, which may have introduced selection bias. Comparative studies across multiple institutions, encompassing both vocational and academic tracks, would provide a more nuanced understanding of contextual and cultural variations in autonomy support and mental health outcomes.

Another limitation concerns the validation strategy. Although the adapted instruments demonstrated acceptable psychometric properties, the measurement model was evaluated within a single sample. This provides weaker evidence than validation based on independent samples or split-sample procedures. Future research should therefore seek to replicate the factor structure in separate samples or use equivalent split-sample methods to strengthen evidence for construct validity (Lorenzo-Seva, 2022).

In addition, a formal back-translation procedure was not implemented during the translation and adaptation of the instruments. Because cross-language adaptation requires careful attention to linguistic and conceptual equivalence, this omission should also be acknowledged as a limitation. Future studies should incorporate more rigorous translation and validation procedures when adapting psychological scales for use in Chinese vocational education

contexts (Kunst et al., 2023; Gouveia et al., 2024). Future research could further investigate additional mediating and moderating mechanisms, particularly basic psychological need satisfaction, perceived competence, relatedness, academic stress, and institutional support, to clarify how autonomy support relates to mental health in vocational college contexts.

6. Conclusion

This study found that perceived autonomy support was positively associated with both cognitive reappraisal and mental health among vocational college students in Shanghai, China. In addition, cognitive reappraisal partially mediated the relationship between perceived autonomy support and mental health, suggesting that autonomy-supportive environments may be linked to student well-being both directly and indirectly through adaptive emotion-regulation processes. These findings highlight the importance of supportive interpersonal environments and emotional skills development in promoting psychological well-being within TVET settings.

The study also offers practical implications for educators, parents, and policymakers. By fostering autonomy-supportive learning and social environments, vocational colleges may help strengthen students' capacity for cognitive reappraisal and support their overall mental health. Overall, the study extends the literature by providing evidence from an underrepresented vocational education population in China and highlights the value of integrating interpersonal support with psychological skill-building in mental health promotion.

Acknowledgement

This study was funded from Universiti Putra Malaysia.

References

- Ahmed, S. P., Ahmed, S. P., Bittencourt-Hewitt, A., & Sebastian, C. L. (2015). Neurocognitive bases of emotion regulation development in adolescence. *Developmental cognitive neuroscience*, 15, 11-25. <https://doi.org/10.1016/j.dcn.2015.07.006>
- Audet, É. C., Levine, S. L., Holding, A. C., Powers, T. A., & Koestner, R. (2021). Peer and family autonomy support during personal goals and crises in emerging adulthood. *Self and Identity*. <https://doi.org/10.1080/15298868.2021.1939772>
- Benita, M. (2020). Freedom to feel: A self-determination theory account of emotion regulation. *Social and Personality Psychology Compass*, e12563. <https://doi.org/10.1111/spc3.12563>
- Black, A. E., & Deci, E. L. (2000). The effects of student self-regulation and instructor autonomy support on learning in a college-level natural science course: A self-determination theory perspective. *Science Education*, 84(6), 740–756.
- Chen, X., & Leung, F. K. (2023). A closer look at teacher support and achievement emotions in Chinese mathematics classrooms: Mediating roles of academic control and intrinsic/extrinsic value. *Educational Psychology*, 43(9), 1084-1101.
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, 41(2), 745–783. <https://doi.org/10.1007/s10490-023-09871-y>
- Chyu, E. P. Y., & Chen, J. K. (2022). The Correlates of Academic Stress in Hong Kong. *International journal of environmental research and public health*, 19(7), 4009. <https://doi.org/10.3390/ijerph19074009>
- Collie, R. J., Martin, A. J., Renshaw, L., & Caldecott-Davis, K. (2024). Students' perceived social-emotional competence: The role of autonomy-support and links with well-being, social-emotional skills, and behaviors. *Learning and Instruction*, 90, 101866. <https://doi.org/10.1016/j.learninstruc.2023.101866>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(7), 1–9. <https://doi.org/10.7275/jyj1-4868>
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Silva, M. M., Teixeira, G., Nunes, E., Lucas, P., & Gaspar, F. (2024). Translation, cross-cultural adaptation, and validation of measurement instruments: A practical guideline for novice researchers. *Journal of Multidisciplinary Healthcare*, 17, 2701–2728. <https://doi.org/10.2147/JMDH.S419714>

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer.

Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.

Deci, E. L., La Guardia, J. G., Moller, A. C., Scheiner, M. J., & Ryan, R. M. (2006). On the benefits of giving as well as receiving autonomy support: Mutuality in close friendships. *Personality and Social Psychology Bulletin*, 32(3), 313–327. <https://doi.org/10.1177/0146167205282148>

Dunning, D. L., Griffiths, K., Kuyken, W., Crane, C., Foulkes, L., & Parker, R. (2019). Research review: The effects of mindfulness-based interventions on cognition and mental health in children and adolescents – a meta-analysis of randomized controlled trials. *Journal of Child Psychology and Psychiatry*, 60(3), 244–258. <https://doi.org/10.1111/jcpp.12980>

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.

Fan, C., & Liu, S. (2024). Exploring the associations among perceived teacher emotional support, resilience, Covid-19 anxiety, and mental well-being: evidence from Chinese vocational college students. *Current Psychology*, 43(16), 14944-14954.

Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage Publications Limited.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.

Gee, D. G., & Cohodes, E. M. (2023). Leveraging the developmental neuroscience of caregiving to promote resilience among youth exposed to adversity. *Development and psychopathology*, 35(5), 2168-2185.

Gouveia, M. do C. L. V., Sveinbjarnardottir, E. K., Rodrigues, M. J. B., Silva, R. M. L. B., Baptista, M. S., & Henriques, M. A. P. (2024). Translation, cultural adaptation, and psychometric validation of the European Portuguese version of the Iceland-Expressive Family Functioning Questionnaire (ICE-EFFQ). *Journal of Family Nursing*, 30(1), 7–29. <https://doi.org/10.1177/10748407231205038>

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2(3), 271–299. <https://doi.org/10.1037/1089-2680.2.3.271>

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage.
- Keyes, C. L. (2009). The mental health continuum: From languishing to flourishing in life (2002). *Foundations of psychological thought: A history of psychology*, 601–617.
- Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, 11(4), 1–10.
- Kong, L. C., & Liu, W. C. (2024). Motivating students using autonomy-supportive instructions: a pilot study. *Asia Pacific Journal of Education*, 1–15. <https://doi.org/10.1080/02188791.2024.2390073>
- Kristensen, S. M., & Jenö, L. M. (2024). The developmental trajectories of teacher autonomy support and adolescent mental well-being and academic stress. *Social Psychology of Education*, 27, 1–32. <https://doi.org/10.1007/s11218-024-09923-1>
- Kunst, J. R., & Bierwiazzonek, K. (2023). Utilizing AI questionnaire translations in cross-cultural and intercultural research: Insights and recommendations. *International Journal of Intercultural Relations*, 97, 101888. <https://doi.org/10.1016/j.ijintrel.2023.101888>
- Lan, X., & Zhang, L. (2019). Shields for Emotional Well-Being in Chinese Adolescents Who Switch Schools: The Role of Teacher Autonomy Support and Grit. *Frontiers in psychology*, 10, 2384. <https://doi.org/10.3389/fpsyg.2019.02384>
- Legate, N., Ryan, R. M., & Weinstein, N. (2012). Is coming out always a “good thing”? Exploring the relations of autonomy support, outness, and wellness for lesbian, gay, and bisexual individuals. *Social Psychological and Personality Science*, 3(2), 145–152. <https://doi.org/10.1177/1948550611411929>
- Li, Q., & Liu, Q. (2023). A meta-analysis of prevalence of mental health problems among higher vocational students in China from 2010 to 2022. *Advances in Psychology*, 13(3), 1256–1268. <https://doi.org/10.12677/AP.2023.133151>
- Lorenzo-Seva, U. (2022). SOLOMON: A method for splitting a sample into equivalent subsamples in factor analysis. *Behavior Research Methods*, 54, 2665–2677. <https://doi.org/10.3758/s13428-021-01750-y>
- Lozano-Jiménez, J. E., Huéscar, E., Moreno-Murcia, J. A., & López-Walle, J. M. (2021). From autonomy support and grit to satisfaction with life through psychological needs and intrinsic motivation of students in higher education. *Frontiers in Psychology*, 11, 579492.
- Ma, C., Ma, Y., & Wang, Y. (2022). Parental Autonomy Support and Mental Health among Chinese Adolescents and Emerging Adults: The Mediating Role of Self-Esteem. *International journal of environmental research and public health*, 19(21), 14029. <https://doi.org/10.3390/ijerph192114029>

Maier, C., Thatcher, J. B., Grover, V., & Dwivedi, Y. K. (2023). Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. *International Journal of Information Management*, 70, 102625. <https://doi.org/10.1016/j.ijinfomgt.2023.102625>

Mammadov, S., Schroeder, N. L., & Heikamp, T. (2023). A meta-analytic review of the relationships between students' perceived teacher and parent autonomy support and positive learning outcomes. *Contemporary Educational Psychology*, 75, 102227.

Marbell-Pierre, K. N., Grolnick, W. S., Stewart, A. L., & Raftery-Helmer, J. N. (2019). Parental autonomy support in two cultures: The moderating effects of adolescents' self-construals. *Child Development*, 90(3), 825–845. <https://doi.org/10.1111/cdev.12947>

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, 12(1), 23–44. <https://doi.org/10.1037/1082-989X.12.1.23>

Meirinhos, A., & Antolín-Suárez, L. (2024). Basic psychological needs and suicidal ideation: testing an integrative model in referred and non-referred adolescents. *Current Psychology*, 43(7), 6438-6449.

Ministry of Education of the People's Republic of China, et al. (2023). *Action plan for strengthening and improving student mental health work in the new era (2023–2025)*. Ministry of Education of the People's Republic of China. Retrieved from http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202305/t20230511_1059224.html

Mohan, L., Roos, L. G., Bui, T. Q., Sichko, S., & Slavich, G. M. (2024). Social support and cognitive reappraisal in adolescent females: The moderating role of lifetime stressor exposure. *International journal of clinical and health psychology: IJCHP*, 24(4), 100525. <https://doi.org/10.1016/j.ijchp.2024.100525>

Morawetz, C., Bode, S., Derntl, B., & Heekeren, H. R. (2017). The effect of strategies, goals and stimulus material on the neural mechanisms of emotion regulation: A meta-analysis of fMRI studies. *Neuroscience & Biobehavioral Reviews*, 72, 111–128. <https://doi.org/10.1016/j.neubiorev.2016.11.014>

Orben, A. (2020). The Sisyphean cycle of technology panics. *Perspectives on Psychological Science*, 15(5), 1143–1157. <https://doi.org/10.1177/1745691620919372>

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>

Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Bloomsbury Publishing PLC.

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, 11, 17–61. <https://doi.org/10.1146/annurev-orgpsych-110721-040030>

Qu, D., Wen, X., Cheng, X., Zhu, A., Wu, Z., Che, L., & Chen, R. (2024). School mental health prevention and intervention strategies in China: a scoping review. *The Lancet regional health. Western Pacific*, 53, 101243. <https://doi.org/10.1016/j.lanwpc.2024.101243>

Qu, Y., & Telzer, E. H. (2017). Cultural differences and similarities in beliefs, practices, and neural mechanisms of emotion regulation. *Cultural diversity & ethnic minority psychology*, 23(1), 36–44. <https://doi.org/10.1037/cdp0000112>

Reeve, J., & Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54–77. <https://doi.org/10.1080/00461520.2020.1862657>

Riepenhausen, A., Wackerhagen, C., Reppmann, Z. C., Dendler, L., Ay, A., Veer, I. M., Voss, G., Golkar, A., & Kalisch, R. (2022). Positive cognitive reappraisal in stress resilience, mental health, and well-being: A comprehensive systematic review. *Emotion Review*, 14(4), 310–331.

Roos, L. G., & Bennett, J. M. (2023). Reappraisal and health: How habitual reappraisal and reappraisal ability interact to protect against life stress in young adults. *Emotion*, 23(5), 1360–1372. <https://doi.org/10.1037/emo0001154>

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>

Ryan, R. M., & Deci, E. L. (2024). Self-determination theory. In *Encyclopedia of quality of life and well-being research* (pp. 6229-6235). Cham: Springer International Publishing.

Sahi, R. S., Eisenberger, N. I., & Silvers, J. A. (2023). Peer facilitation of emotion regulation in adolescence. *Developmental Cognitive Neuroscience*, 62, 101262. <https://doi.org/10.1016/j.dcn.2023.101262>

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74. <https://doi.org/10.23668/psycharchives.12784>

Shen, J., Jia, J., Wang, L., & Fang, X. (2023). Autonomy-relatedness patterns and their association with academic and psychological adjustment among Chinese adolescents. *Journal of youth and adolescence*, 52(6), 1272-1286.

Shum, C., Dockray, S., & McMahon, J. (2025). The relationship between cognitive reappraisal and psychological well-being during early adolescence: A scoping review. *Journal of Early Adolescence*, 45(1), 104–133. <https://doi.org/10.1177/02724316241231918>

Song, H., Kwon, M., & Kim, H. S. (2024). Differences and similarities in the use of nine emotion regulation strategies between Western and East-Asian individuals: A systematic review and meta-analysis. *Journal of Cross-Cultural Psychology*. <https://doi.org/10.1177/00220221241285006>

Stover, A. D., DeJoseph, M. L., Jazaieri, H., & Forbes, C. E. (2024). A meta-analysis of cognitive reappraisal and personal resilience. *Clinical Psychology Review*, 111, 102434.

Su, Y.-L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review*, 23(1), 159–188. <https://doi.org/10.1007/s10648-010-9142-7>

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

The State Council of the People's Republic of China. (2010). *Outline of China's national plan for medium- and long term education reform and development (2010–2020)*. Government of the People's Republic of China. Retrieved from https://www.moe.gov.cn/srcsite/A01/s7048/201007/t20100729_171904.html

Tilga, H., Kalajas-Tilga, H., Hein, V., Raudsepp, L., & Koka, A. (2021). Perceived autonomy support from peers, parents, and physical education teachers as predictors of physical activity and health-related quality of life among adolescents: A one-year longitudinal study. *Education Sciences*, 11(9), 457.

Troy, A. S., Wilhelm, F. H., Shallcross, A. J., & Mauss, I. B. (2010). Seeing the silver lining: cognitive reappraisal ability moderates the relationship between stress and depressive symptoms. *Emotion (Washington, D.C.)*, 10(6), 783–795. <https://doi.org/10.1037/a0020262>

Van der Kaap-Deeder, J., Soenens, B., Mouratidis, A., De Pauw, S., Krøjgaard, P., & Vansteenkiste, M. (2020). Towards a detailed understanding of preschool children's memory-related functioning and emotion regulation: The role of parents' observed reminiscence style, memory valence, and parental gender. *Developmental Psychology*, 56(9), 1696–1708. <https://doi.org/10.1037/dev0001048>

Vasconcellos, D., Parker, P. D., Hilland, T., Cinelli, R., Owen, K. B., Kapsal, N., ... & Lonsdale, C. (2020). Self-determination theory applied to physical education: A systematic review and meta-analysis. *Journal of educational psychology*, 112(7), 1444.

Wei, M., Su, J. C., Carrera, S., Lin, S.-P., & Yi, F. (2013). Suppression and interpersonal harmony: A cross-cultural comparison between Chinese and European Americans. *Journal of Counseling Psychology*, 60(4), 625–633. <https://doi.org/10.1037/a0033413>

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: a test of self-determination theory. *Journal of personality and social psychology*, 70(4), 767.

Yang, D., Chen, P., Wang, H., Wang, K., & Huang, R. (2022). Teachers' autonomy support and student engagement: A systematic literature review of longitudinal studies. *Frontiers in Psychology*, 13, 925955. <https://doi.org/10.3389/fpsyg.2022.925955>

Ye, B., Zhao, S., Zeng, Y., Chen, C., & Zhang, Y. (2022). Perceived parental support and college students' depressive symptoms during the COVID-19 pandemic: The mediating roles of emotion regulation strategies and resilience. *Current Psychology*. Advance online publication. <https://doi.org/10.1007/s12144-022-03049-3>

Zachariou, A., & Bonneville-Roussy, A. (2024). The role of autonomy support from teachers in young learners' self-regulation in dyadic contexts: An examination through three-level multilevel analysis. *Learning and Instruction*, 89, 101843. <https://doi.org/10.1016/j.learninstruc.2023.101843>

Zeng, Y., Wang, G., Xie, C., Hu, X., & Reinhardt, J. D. (2019). Prevalence and correlates of depression, anxiety and symptoms of stress in vocational college nursing students from Sichuan, China: a cross-sectional study. *Psychology, health & medicine*, 24(7), 798-811.

Zhao, J., & Qin, Y. (2021). Perceived teacher autonomy support and students' deep learning: The mediating role of self-efficacy and the moderating role of perceived peer support. *Frontiers in Psychology*, 12, 652796. <https://doi.org/10.3389/fpsyg.2021.652796>

Zhou, L. H., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2019). Effects of perceived autonomy support from social agents on motivation and engagement of Chinese primary school students: Psychological need satisfaction as mediator. *Contemporary Educational Psychology*, 58, 323–330.

Appendix

Part A: Mental Health Continuum- Short Form (MHC-SF)

Items	Code	Factor Loading
1. happy	EWB 1	0.611
2. interested in life	EWB 2	0.631
3. satisfied with life	EWB 3	0.609
4. that you had something important to contribute to society	SWB 1	0.659
5. that you belonged to a community (like a social group, your school, or your neighborhood)	SWB 2	0.699
6. that our society is a good place, or is becoming a better place, for all people	SWB 3	0.682
7. that people are basically good	SWB 4	0.648
8. that the way our society works made sense to you	SWB 5	0.671
9. that you liked most parts of your personality	PWB 1	0.653
10. good at managing the responsibilities of your daily life	PWB 2	0.710
11. that you had warm and trusting relationships with others	PWB 3	0.657
12. that you had experiences that challenged you to grow and become a better person	PWB 4	0.769
13. confident to think or express your own ideas and opinions	PWB 5	0.632
14. that your life has a sense of direction or meaning to it	PWB 6	0.726

Part B: Emotional Regulation Questionnaire (ERQ)

Items	Code	Factor Loading
1.When I want to feel more positive emotion (such as joy or amusement), I change what I'm thinking about.	ERQ 1	0.763
2.When I want to feel less negative emotion (such as sadness or anger), I change what I am thinking about.	ERQ 2	0.823
3.When I am faced with a stressful situation, I make myself think about it in a way that helps me stay calm.	ERQ 3	0.813
4.When I want to feel more positive emotion, I change the way I'm thinking about the situation.	ERQ 4	0.850
5.I control my emotions by changing the way I think about the situation I am in.	ERQ 5	0.709
6.When I want to feel less negative emotion, I change the way I am thinking about the situation.	ERQ 6	0.785

Part C: Autonomy Support Questionnaire (ASQ)

Items	Code	Factor Loading
1.I feel that my family provides me with choices and options.	ASQP 1	0.704
2.I believe that my family is very understanding of me.	ASQP 2	0.612
3.My family conveys confidence in my abilities.	ASQP 3	0.570
4.I feel that my family accepts me.	ASQP 4	0.661
5.I believe that my family trusts me.	ASQP 5	0.663
6.My family listens to my thoughts and ideas.	ASQP 6	0.714
7.My family encourages me to express my true emotions.	ASQP 7	0.717
8.I feel that my family cares about me as a person.	ASQP 8	0.661
9.My family tries to understand how I see things.	ASQP 9	0.675
10.I feel able to share my feelings with my family.	ASQP 10	0.588
Perceived Friends Autonomy Support		
1.I feel that my friends provide me with choices and options.	ASQF 1	0.630
2. I believe that my friends are very understanding of me.	ASQF 2	0.615
3. My friends convey confidence in my abilities.	ASQF 3	0.623
4. I feel that my friends accept me.	ASQF 4	0.641
5. I believe that my friends trust me.	ASQF 5	0.768
6. My friends listen to my thoughts and ideas.	ASQF 6	0.766
7. My friends encourage me to express my true emotions.	ASQF 7	0.655
8. I feel that my friends care about me as a person.	ASQF 8	0.685
9. My friends try to understand how I see things.	ASQF 9	0.679
10. I feel able to share my feelings with my friends.	ASQF 10	0.695
Perceived School Peers Autonomy Support		
1.I feel that my school peers provide me with choices and options.	ASQSP 1	0.595
2.I believe that my school peers are very understanding of me.	ASQSP 2	0.711
3.My school peers convey confidence in my abilities.	ASQSP 3	0.626
4.I feel that my school peers accept me.	ASQSP 4	0.606
5.I believe that my school peers trust me.	ASQSP 5	0.623
6.My school peers listen to my thoughts and ideas.	ASQSP 6	0.540
7.My school peers encourage me to express my true emotions.	ASQSP 7	0.646
8.I feel that my school peers care about me as a person.	ASQSP 8	0.665
9.My school peers try to understand how I see things.	ASQSP 9	0.650
10.I feel able to share my feelings with my school peers.	ASQSP 10	0.651

Part D: The Learning Climate Questionnaire (LCQ)

Items	Code	Factor Loading
1.I feel that my teacher provides me choices and options.	LCQ 1	0.708
2.I feel understood by my teacher.	LCQ 2	0.776
3.I am able to be open with my teacher during class.	LCQ 3	0.737
4.My teacher conveyed confidence in my ability to do well in the course.	LCQ 4	0.803
5.I feel that my teacher accepts me.	LCQ 5	0.778
6.My teacher made sure I really understood the goals of the course and what I need to do.	LCQ 6	0.723
7.My teacher encouraged me to ask questions.	LCQ 7	0.727
8.I feel a lot of trust in my teacher.	LCQ 8	0.716
9.My teacher answers my questions fully and carefully.	LCQ 9	0.800
10.My teacher listens to how I would like to do things.	LCQ 10	0.750
11.My teacher handles people's emotions very well.	LCQ 11	0.736
12.I feel that my teacher cares about me as a person.	LCQ 12	0.754
13.I don't feel very good about the way my teacher talks to me. *	LCQ 13	0.584
14.My teacher tries to understand how I see things before suggesting a new way to do things.	LCQ 14	0.750
15.I feel able to share my feelings with my teacher.	LCQ 15	0.712

The Manifestation and Implications of Academic Entitlement Among University Students in Henan Province, China

ABSTRACT

Authors: Guizhi Song¹, Muhammad Asyraf Che Amat^{1*} & Maizura Yasin¹

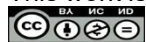
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:

mhdasyraf@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Academic entitlement refers to a student's subjective belief that they are deserving of academic success even without personal effort or achievement. This phenomenon has been increasingly discussed in higher education literature, with researchers noting a growing trend of entitled attitudes and demanding behaviors among university students in Europe and America across various domains. However, research on academic entitlement within the Chinese student population is limited. Consequently, this study investigates the current state of academic entitlement among university students in Henan Province within the Chinese cultural context and discusses its implications for future research. Utilizing a quantitative research method, the study sampled 886 university students from Henan Province, China, for analysis. The results showed that the academic entitlement level of college students, represented by those from Henan Province, was at a moderately low level, with only 33.4% of students scoring more than 3 points on the academic entitlement scale, indicating that academic entitlement may not be a common phenomenon among college students in Henan Province. Additionally, the study showed that, apart from fourth-year students, whose academic entitlement scores were significantly lower than those of other grade levels, no significant differences in academic entitlement were observed across other demographic variables (gender, number of children, birth order, academic discipline). These findings contrast with those from European and American studies, suggesting that academic entitlement may be associated with cultural context. This study offers three implications for future research (see Discussion): longitudinal designs to examine temporal changes, qualitative or case study methods to explore typical manifestations, and expanded sampling to test generalizability.

Keywords: Academic Entitlement, University Students, Chinese Cultural Context, Henan Province.

1. Introduction

Academic Entitlement (AE) reflects a diminished sense of personal responsibility in academic tasks, a belief that one deserves rewards regardless of effort, and irrational expectations for unearned accommodations or grades (Carollo, 2020; Yu & Li, 2020). It is generally conceptualized as a set of beliefs and behaviors with negative connotations. Since its inception, the concept of AE has garnered sustained academic attention. As early as 2008, Greenberger et al. (2008) noted that the widespread sense of entitlement had attracted significant attention from both the media and researchers. In recent years, AE has become a topic of considerable debate within academia (Knepp & Knepp, 2022).

1.1 The Universality of Academic Entitlement

Currently, academic entitlement is regarded by university educators and administrators as a major concern within higher education. Both researchers and practitioners have observed a growing trend among college students characterized by an increased sense of entitlement and demanding attitudes, a phenomenon noted across various domains of educational research (Halat et al., 2024). This perspective has been substantiated by a considerable body of empirical evidence.

A study by Canadian scholar Huang (2017) found that 80% of participants endorsed items on an entitlement scale. Finney and Finney (2010) reported that 58% of individuals aged 18–29 believed they were entitled to privileges. Since extra credit is often a significant factor in achieving higher grades, the expectation of such opportunities has become a typical manifestation of AE. Hassel and Lourey (2005) found that 62% of students believed it was the instructor's responsibility to offer extra credit. Similarly, Schaefer et al. (2013) observed that more than half (55%) of university students indicated that instructors should provide extra credit assignments to help students improve their grades. Zare (2022) further confirmed that nearly all instructors had been asked by students to provide extra credit. These data illustrate that AE is a highly prevalent phenomenon among university students.

1.2 The Negative Effects of Academic Entitlement

Academic Entitlement exerts multiple negative effects, primarily manifested in four aspects:

First, AE adversely impacts students' academic performance. Kurtyilmaz (2019) identified AE as one of the most critical factors affecting student achievement, noting a strong correlation between AE and poor academic outcomes. Students with high levels of AE are more prone to class absenteeism, using smartphones during lectures, and showing disrespect toward peers and instructors (Lippmann et al., 2009; Mellor, 2011). They are also more likely to exhibit hostile, controlling, and aggressive behaviors—including challenging peer relationships, intentional harm, greed, and overt aggression (Abdellatif, 2022). Furthermore, AE serves as a significant predictor of academic misconduct, including plagiarism, cheating, and academic fraud (Stiles et al., 2018; Chapin, 2018; Vallade & Martin, 2014). These behaviors are bound to lead to poor academic performance among students.

Second, AE impairs students' psychological well-being. As a maladaptive belief and behavioral tendency, AE inevitably affects mental health and may contribute to the development of pathological traits (Boswell, 2012). It is well established that low self-esteem is associated with

various psychological issues. Greenberger et al. (2008) found that students with higher AE tend to have lower self-esteem. Individuals with low self-esteem are more susceptible to negative emotions such as anxiety, depression, and loneliness.

Third, AE negatively influences students' subjective experiences and interpersonal relationships. Students with high AE often believe they are entitled to make excessive demands of instructors and administrators. When these expectations are not met, they frequently experience frustration, dissatisfaction, anger, or hostility. Moreover, due to a lack of regard for how their actions affect others, highly entitled students are more likely to attempt to control others to satisfy their own needs or vent dissatisfaction through complaints and confrontations (Yu & Li, 2020). As a result, they often experience poor interpersonal relationships.

Fourth, AE exacerbates inequity in the academic environment. Students with high AE typically perform poorly academically yet believe they deserve high grades. They may negotiate with instructors to obtain undeserved scores, which inevitably creates an unfair academic environment, particularly disadvantaging diligent and sincere students. Additionally, AE contributes to grade inflation. Chowdhury (2018) argued that grade inflation has become a widespread phenomenon in universities and colleges globally, with nearly all higher education institutions relying on student evaluations for faculty assessment. Consequently, some students engage in transactional behaviors such as bargaining for higher grades in exchange for positive course evaluations (Gruhlke, 2018).

Given the increasing concern surrounding academic entitlement in higher education, the present study aims to investigate AE among Chinese university students. This research addresses two questions:

First, what is the level of AE among university students in Henan Province?

Second, are there differences in AE based on gender, grade, number of children, birth order, or academic discipline? Given China's one-child policy, which has been in effect for decades, a considerable proportion of current university students are only children. This unique demographic context makes family size (i.e., number of children) a potentially relevant variable for understanding academic entitlement in the Chinese setting, in addition to conventional variables such as gender and grade level. Beyond these empirical questions, the study also seeks to interpret the findings within the Chinese cultural context, with particular attention to how cultural norms and educational practices may shape the manifestation of academic entitlement.

This study makes several contributions to the existing literature on Academic Entitlement. First, it provides evidence from a large mainland Chinese sample drawn from Henan Province, a region that is demographically representative yet underexplored in prior AE research. Second, it replicates and extends the psychometric validation of the Academic Entitlement Questionnaire (AEQ) in the Chinese context, building upon the work of Hu et al. (2021). Third, it examines demographic correlates of AE—such as grade level, birth order, and only-child status—that have received limited attention in mainland samples. Collectively, these contributions offer a more contextually grounded understanding of AE and help clarify whether previously observed patterns generalize beyond Western and small-scale Asian samples.

2. Literature Review

Research on Academic Entitlement has predominantly been conducted in European and American countries. A review of the literature using "academic entitlement" as a keyword reveals that the majority of articles draw their samples from these regions. Due to the relatively mature higher education systems and pronounced student rights consciousness in Western societies, AE has attracted academic attention there at an earlier stage. This has led to the widespread yet unsubstantiated assumption that AE is a common and universal phenomenon, even though most supporting evidence comes from students in American institutions (Blincoe & Garriss, 2017). Blincoe and Garriss (2017) regard this as a noteworthy issue in the existing literature. In reality, there remains a scarcity of studies exploring AE in non-Western contexts. This does not imply that AE is absent among university students in these regions; rather, it highlights the need for more empirical research to verify its prevalence and manifestations. Many researchers agree that AE is not exclusive to individualistic Western cultures but should be regarded as a shared concern for educators in both Eastern and Western contexts (Huang & Kuo, 2020; Blincoe & Garriss, 2017).

Due to differences in socialization processes and value systems, the manifestations of AE may vary significantly across cultures. In Western individualistic cultures, AE often takes the form of externalizing responsibility, asserting one's rights, and directly challenging institutional rules (Rao & Kay, 2026). In contrast, within the Chinese collectivist society rooted in Confucian culture, where emphasis is placed on diligent study, respect for authority, and self-discipline (Lu et al., 2023), students are less likely to openly question their instructors or explicitly externalize responsibility. Hong et al. (2016) found that individualism was significantly associated with AE, whereas collectivism showed no significant correlation with AE. Similarly, Ellis et al. (2021) suggested that while AE is prevalent among Asian students, it tends to manifest in more implicit ways. Therefore, in the context of mainland China, the dimensions of AE may exhibit different patterns of intensity and mechanisms of action compared to those observed in Western settings. Based on this cultural perspective, this study aims to explore the occurrence of AE among Chinese university students, represented by those from Henan Province, within Chinese cultural context.

First, a number of studies confirm the cross-cultural relevance of AE research. Hong et al. (2012) investigated the relationship between relevant variables and AE among students from eight universities in Taiwan, confirming the presence of AE among university students in the region. Gwinner et al. (2021) surveyed business students in Taiwan and found that AE was not only prevalent but also manifested at relatively high levels. Ellis et al. (2021), in two consecutive studies, found a small yet statistically significant increase in AE among Asian students compared to their White counterparts. Furthermore, research in Turkey has explored the applicability of the Academic Entitlement Questionnaire (AEQ), developed by Kopp et al., within Turkish culture (Kurtyilmaz, 2019). Scholars in countries and regions such as Saudi Arabia have also conducted studies on the relationship between AE and professional perception (Halat et al., 2024) as well as between AE and moral intelligence (Al-Adamat et al., 2024). These studies collectively demonstrate both the feasibility and necessity of cross-cultural research on AE.

Furthermore, Chinese university students also exhibit many behaviors associated with AE. Disruptive classroom conduct, disrespect towards instructors, and cheating have been empirically linked to AE, and such behaviors are relatively widespread among students in China. Wu (2019) found that university instructors in China have reported an increase in disruptive classroom behavior in recent years, noting a lack of self-discipline among students. Moreover, the widespread implementation of “student evaluations of teaching” in many Chinese institutions has led to a phenomenon where instructors inflate grades to secure higher student ratings. This “implicit collusion” between students and instructors contributes to a decline in teaching quality (Lu & Jian, 2021) and may also foster increased AE among students. Additionally, academic cheating is not uncommon in Chinese higher education. A study by Lin and Pu (2018) surveying four universities in China revealed that 12.2% of students admitted to having engaged in cheating on exams. These behaviors have all been empirically linked to AE.

Finally, research focusing on mainland Chinese university students remains insufficient. Hu et al. (2021) adapted the Academic Entitlement Questionnaire (AEQ), originally developed by Kopp et al., into Chinese and examined its reliability and validity, thereby establishing a foundational tool for AE research in China. In the same year, Lu and Jian (2021) explored the influence of student background, course learning experience, and academic self-efficacy on AE. Shafait and Sahibzada (2024) collected data from higher education institutions in Xi'an, Sichuan, and Hangzhou, revealing a positive and significant direct relationship between the Dark Triad (DT) personality traits and AE. Additionally, several studies have investigated AE among nursing students. One research team focused on the impact of parental overprotection on AE in nursing students, finding significant correlations among all variables and identifying external locus of control and psychological entitlement as sequential mediators between parental overprotection and academic entitlement (Peng et al., 2022; Peng et al., 2024). Sang et al. (2023) examined the status and influencing factors of AE among undergraduate nursing students, indicating that AE levels were relatively high, with career choice autonomy, family affection index, and learning self-efficacy serving as significant influencing factors. Zhou (2023) compared AE between undergraduate and diploma nursing students, finding that diploma students exhibited higher levels of AE than undergraduate students ($p < 0.05$). Overall, there is still a scarcity of studies targeting students in general universities and colleges, underscoring the necessity for further research on AE in the Chinese context.

The present study adopts a theoretically informed yet non-causal framework to contextualize academic entitlement and its variation across student groups. Drawing on Social Learning Theory, AE can be understood as a belief system developed through repeated social interactions, observational learning, and reinforcement processes (Bandura, 1977). In academic settings, students may internalize entitlement-related beliefs when they observe that outcomes such as grades or academic accommodations are not consistently contingent upon effort, but can instead be influenced through negotiation, institutional practices, or external support.

At the same time, ecological systems theory (Bronfenbrenner, 2000) suggests that individual differences in such beliefs are shaped by variations in developmental contexts. Demographic characteristics—including gender, grade level, birth order, and only-child status—can be viewed as proxies for differences in socialization experiences, family expectations, and educational environments. These contextual variations may lead to differential exposure to reinforcement patterns and expectations, thereby contributing to variability in AE. Taken

together, these perspectives provide a conceptual foundation for examining not only the overall level of AE but also its distribution across demographic groups.

3. Methodology

The main purpose of this study is to explore the AE level and performance of Chinese college students from Henan Province, and to examine the impact of demographic variables (gender, grade, number of children, birth order, academic discipline) on AE. The participants in this study were full-time undergraduate students, and the study has been approved by the ethics review department of Universiti Putra Malaysia (UPM), approval number JKEUPM-2024-582.

3.1. Research Tools

This study employed a quantitative research design and utilized the revised Chinese version of the Academic Entitlement Questionnaire (AEQ) developed by Hu et al. (2021). The original AEQ, developed by Kopp et al. (2011), conceptualizes academic entitlement as a unidimensional construct reflecting generalized entitlement beliefs. Although some studies have proposed multidimensional conceptualizations of AE, the AEQ remains one of the most widely used instruments due to its parsimony and empirical support (Kinne et al., 2022).

Hu et al. (2021) translated, back-translated, and culturally adapted the AEQ for use in the Chinese context. The resulting scale retained the original unidimensional structure, consisting of eight items. Their validation study, based on a sample of 1,102 university students in Guizhou Province, demonstrated acceptable internal consistency (Cronbach's $\alpha = 0.76$) and satisfactory test-retest reliability ($r = 0.70$, $p < .01$). In addition, confirmatory factor analysis (CFA) indicated a good model fit (CFI = 0.93, TLI = 0.90, IFI = 0.93, GFI = 0.97, NFI = 0.91, RMSEA = 0.07, SRMR = 0.04), supporting the structural validity of the instrument.

However, given ongoing debates regarding the dimensionality of AE and the potential influence of cultural context on its measurement, further validation across different Mainland Chinese samples is warranted. Therefore, the present study adopts the Chinese AEQ and conducts CFA to re-examine its factor structure in a larger and regionally distinct sample from Henan Province. The detailed items of the scale can be found in Appendix 1 (Chinese and English).

3.2. Sampling Methods and Sample Size

This study employed a stratified cluster sampling method to conduct an on-site questionnaire survey among full-time four-year undergraduate students at a university in Xinxiang City, Henan Province, China. Data collection was carried out during the regular teaching period in November and December 2024.

First, one major from the liberal arts and one major from the sciences were randomly selected from the university. After obtaining permission from department administrators and faculty members, questionnaires were administered on site during breaks between classes. All participants provided informed consent prior to participation and completed the survey voluntarily.

A total of 950 questionnaires were distributed, of which 907 were returned, yielding a response rate of 95.47%. After data screening, 21 questionnaires were excluded due to incomplete responses or identical and systematically patterned answers across items. The final valid sample consisted of 886 participants, corresponding to an effective rate of 97.68%.

To justify the adequacy of the sample size, an a priori power analysis was conducted using G*Power 3.1. A small effect size ($f = 0.15$) was assumed based on the conventional benchmarks proposed by Jacob Cohen (1988), which are widely adopted in social and educational research. With an alpha level of .05 and a desired statistical power of .80 for one-way ANOVA, the required minimum sample size was 344. It was substantially smaller than the actual sample size ($N = 886$), indicating that the present study was sufficiently powered to detect small effects. Notably, the observed effect size for grade differences ($\eta^2 = 0.02$; $f \approx 0.14$) closely aligns with the assumed small effect size, further supporting the adequacy of the sample size.

3.3. Test of Validity and Reliability

To examine whether common method variance might influence the results, an exploratory factor analysis was performed in which all measurement items were entered simultaneously. The analysis yielded two factors with eigenvalues greater than one. Importantly, the first unrotated factor explained 37.65% of the total variance, which did not exceed the commonly accepted threshold of 40% (Fuller et al., 2016). These results indicate that common method bias is unlikely to substantially affect the validity of the present findings.

The internal consistency of the AEQ was satisfactory, as reflected by a Cronbach's α value of 0.74. To further investigate the construct validity of the scale, exploratory factor analysis (EFA) was conducted. The Kaiser–Meyer–Olkin (KMO) index was 0.82, and Bartlett's test of sphericity reached statistical significance, confirming the suitability of the data for factor analysis.

Table 1. Communalities of Items from the Exploratory Factor Analysis

	Initial	Extraction
AE1	0.23	0.25
AE2	0.01	0.003
AE3	0.21	0.22
AE4	0.21	0.23
AE5	0.36	0.43
AE6	0.35	0.38
AE7	0.31	0.37
AE8	0.40	0.51

Note. Initial communalities represent the proportion of variance explained by all factors prior to extraction. Extracted communalities represent the proportion of variance explained by the retained factor(s).

As shown in Table 1, exploratory factor analysis (EFA) was conducted using principal axis factoring (PAF) with a single-factor solution specified, consistent with the theoretical assumption of a unidimensional AE construct. The results indicated that communalities varied

across items. Notably, AE2 exhibited an extremely low communality (0.003), suggesting that it shared minimal variance with the underlying factor. Consistent with this, communalities were low for several items: AE1 (0.25), AE2 (0.003), AE3 (0.22), AE4 (0.23). While the original AEQ has demonstrated acceptable psychometric properties in Western samples (Kopp et al., 2011), the present results suggest that several items—particularly AE2 (“Professors should only lecture on material covered in the textbook and assigned readings”)—may function poorly as indicators of academic entitlement among university students in mainland China. This could reflect genuine cultural non-equivalence of certain entitlement-related beliefs. In the Chinese educational context, where lectures often go beyond textbooks and students are generally expected to follow the instructor’s lead, endorsing or rejecting this item may carry different meanings than in Western settings. Thus, the low variance explained and low communalities may not merely reflect sampling variation but rather point to the need for cultural adaptation or revision of the AEQ for use in Chinese student populations. Future research should consider developing or validating context-specific measures of academic entitlement that better capture its manifestations in non-Western educational environments.

To further evaluate the measurement structure, confirmatory factor analysis (CFA) was conducted using AMOS 26.0. Two models were specified and compared: the original 8-item model and a revised 7-item model that excluded AE2. As shown in Appendix 2, the 8-item model yielded mixed fit indices (CFI = .91, TLI = .88, RMSEA = .08), although the χ^2/df ratio was high (6.93). The 7-item model did not improve model fit; instead, fit indices deteriorated slightly (CFI = .91, TLI = .87, RMSEA = .10, χ^2/df = 9.65). The changes in fit indices (ΔCFI = $-.002$; ΔRMSEA = $+.017$) indicate that removing AE2 did not meaningfully enhance model fit. SRMR was available for the 7-item model (.06) but was not reported for the 8-item model in the AMOS output. The absence of SRMR for the 8-item model is due to software output limitations rather than analytical omission. Therefore, model evaluation primarily relied on other widely recommended fit indices, including CFI, TLI, RMSEA, and χ^2/df . In addition, modification indices were examined to identify potential areas for model improvement; however, no post hoc modifications were implemented due to the lack of clear theoretical justification.

Table 2. Standardized Factor Loadings from Confirmatory Factor Analysis

Item	Loading (Std.)	S.E.	C.R.	p
AE1	0.48	—	—	—
AE2	0.06	0.05	1.51	0.13
AE3	0.44	0.08	9.80	<.001
AE4	0.46	0.09	10.01	<.001
AE5	0.66	0.10	11.72	<.001
AE6	0.64	0.09	11.42	<.001
AE7	0.62	0.11	11.55	<.001
AE8	0.72	0.10	12.53	<.001

Note. Values represent standardized factor loadings obtained from confirmatory factor analysis.

As reported in Table 2, the standardized factor loading of AE2 was extremely low (.06), whereas the remaining items showed moderate to strong loadings (.44–.72). This result is consistent with the EFA findings and further indicates that AE2 contributes minimally to the latent construct. Composite reliability (CR) and average variance extracted (AVE) were

calculated to assess reliability and convergent validity. For the 8-item model, CR was 0.75, exceeding the recommended threshold of 0.70, whereas AVE was 0.30, below the suggested criterion of 0.50 (Hair et al., 2010). After removing AE2, CR increased slightly to 0.78, while AVE improved to 0.34. Although the AVE values were below the recommended threshold of 0.50, this result can be attributed to the relatively moderate factor loadings of several items and the presence of an item (AE2) with extremely low loading. In constructs reflecting broad psychological attitudes, it is not uncommon for AVE to fall below the suggested criterion, particularly when multiple indicators capture diverse aspects of the construct.

Moreover, prior research has suggested that adequate composite reliability may compensate for lower AVE values (Fornell & Larcker, 1981). Given that CR values exceeded the recommended threshold and all retained items demonstrated significant factor loadings, the convergent validity of the construct below conventional thresholds, but CR was adequate. This pattern is commonly observed in studies involving attitudinal or perceptual constructs. Although AE2 demonstrated extremely low communality (0.003) and a near-zero factor loading (0.06), suggesting that this item does not capture academic entitlement as understood in the Chinese cultural context, its removal did not improve overall model fit. Nevertheless, on conceptual grounds—given that the belief “professors should only lecture on the textbook” is rarely endorsed or meaningfully interpretable in Chinese educational settings where lectures routinely go beyond prescribed materials—AE2 was excluded from subsequent analyses. The resulting 7-item scale showed improved content relevance for the present sample, though model fit remained suboptimal ($\chi^2/df = 9.65$; RMSEA = 0.10). These results highlight the need for cultural adaptation of the AEQ rather than indicating a true absence of academic entitlement among Chinese students.

4. Findings and Discussion

Data processing and statistical analysis were performed using the SPSS 29.0 software package. To address the three research objectives, the means and standard deviations of the seven items on the AEQ were compared. Independent-samples t-tests and one-way ANOVA were conducted to examine the effects of five demographic variables on AE. Basic information regarding the demographic variables is presented in Table 3.

Table 3. Demographic Factors of the Respondents

Demographic Variable	Category	Frequency	Percentage (%)
Gender	Male	383	43.2
	Female	503	56.8
Grade	Freshman	396	44.7
	Sophomore	247	27.9
	Junior	115	13.0
	Senior	128	14.2
Number of children	Only children	88	9.9
	Non-only children	798	90.1
Birth order	First	435	49.1
	Second	377	42.6
	Third	61	6.9
	Fourth	13	1.5

Academic discipline	Liberal Arts	325	36.7
	Sciences	561	63.3

Note. N = 886.

It should be specifically noted that in this study, only 88 out of the surveyed college students were from single-child families, accounting for 9.9%. The non-only-child college students numbered 798, accounting for 90.1%. This indicates that among the sample population of this study, non-only-child college students are in the majority. Furthermore, in terms of birth order, 61 college students ranked third, accounting for 6.9%, while only 13 college students ranked fourth in their families, accounting for 1.5%. This indicates that the majority of Chinese families have 1 or 2 children.

4.1. Average Levels and Performance of Academic Entitlement

Table 4 presents the descriptive statistics, including the minimum, maximum, mean, and standard deviation, for each item.

Table 4. Descriptive Statistics for All Items

	N	Minimum	Maximum	Mean	SD
AE1	886	1.00	5.00	3.05	1.13
AE3	886	1.00	5.00	3.02	1.02
AE4	886	1.00	5.00	3.07	1.12
AE5	886	1.00	5.00	2.21	0.97
AE6	886	1.00	5.00	2.18	0.88
AE7	886	1.00	5.00	2.72	1.12
AE8	886	1.00	5.00	2.21	0.93
AE	886	1.00	5.00	2.64	0.67

As shown in Table 4, the overall level of academic entitlement among Chinese university students is moderately low, with a mean score of 2.64 (SD = 0.67). AE1 (“If I don’t do well on a test, the professor should make tests easier or curve the grades”), AE3 (“If I am struggling in a class, the professor should approach me and offer to help”), and AE4 (“It is the professor’s responsibility to make it easy for me to succeed”) (Kopp et al., 2011) received scores slightly above 3. AE1 corresponds to the “entitlement to knowledge” dimension, while AE 3 and AE4 reflect “externalization of responsibility”(Kopp et al., 2011). These survey results indicate that students may believe they have the right to acquire knowledge, and that professors should be held accountable for their academic performance; however, the scores for these three items were only slightly above 3 points, and were still far from reaching the 4 points.

The remaining four items all scored below 3. Among them, AE5 and AE6, which represent “external attribution of learning problems”(Kopp et al., 2011) also received relatively low scores, suggesting that Chinese students are not inclined to attribute all learning difficulties to external factors. AE7 (“No matter the reason for my absence, I should be allowed to make up the exam”) reflects “student control over course policies” (Kopp et al., 2011) and AE8 (“Because I paid tuition, I deserve a passing grade”) corresponds to “entitled expectations of

outcomes”(Kopp et al., 2011). These items received low scores, indicating that Chinese college students may not fully agree with most of the items on the AEQ. One possible explanation for the relatively lower scores on externalisation-related items is the influence of family education norms within the Chinese cultural context. Chinese families typically emphasise effort, responsibility, and respect for authority in educational settings. As a result, students may be less likely to attribute academic outcomes to external factors such as teachers or institutions, and instead show a stronger tendency toward internal responsibility. This pattern is consistent with prior studies highlighting the role of collectivist values and effort-based beliefs in shaping students’ academic attitudes. In order to better understand how many college students scored 3 points or higher, this study conducted a statistical analysis of the distribution of scores within different score ranges and the corresponding percentages. The analysis results are presented in Table 5.

Table 5: Distribution and Percentages of Academic Entitlement Scores

Scores Range	n	Percent(%)	Cumulative Percent(%)
5.00	4	0.50	0.50
4.00–4.99	22	2.40	2.90
3.00–3.99	270	30.50	33.40
2.00–2.99	475	53.60	87.00
1.00–1.99	115	13.00	100.00

Note. Scores represent the mean academic entitlement score (range = 1–5). N = 886

Descriptive analysis indicated that 296 students (33.4%) obtained a mean AE score of 3.0 or above, suggesting that approximately one-third of the sample demonstrated moderate to relatively high levels of academic entitlement. To examine whether the prevalence of AE was associated with demographic characteristics, a binary logistic regression analysis was conducted. AE was dichotomized as ≥ 3.0 versus < 3.0 , with gender, grade, and number of children entered simultaneously into the model. The results indicated that none of these variables significantly predicted the likelihood of scoring 3.0 or above (all $ps > .05$), suggesting that AE prevalence was not strongly associated with these demographic factors.

Table 6. Binary Logistic Regression Predicting Academic Entitlement (≥ 3.0 vs. < 3.0)

	B	S.E.	Wald	p	OR	95% C.I. for OR
Gender(1)	-0.22	0.15	2.27	0.13	0.80	[0.60, 1.07]
Grade	-0.09	0.07	1.84	0.18	0.91	[0.79, 1.04]
Number of children(1)	0.02	0.24	0.01	0.93	1.02	[0.64, 1.62]

Note. OR = odds ratio; CI = confidence interval; Gender was coded as 1 = male and 2 = female; Number of children was coded as 1 = only child and 2 = non-only-child.

In the present study, academic entitlement was dichotomized using a cut-off point ($AE \geq 3.0$) to facilitate logistic regression analysis. However, it is important to note that dichotomizing continuous variables may lead to a loss of information and reduced statistical power. As noted by MacCallum et al. (2002), the practice of dichotomization can obscure underlying variability and potentially distort relationships among variables. Therefore, to address this limitation, the present study also conducted analyses using the continuous AE variable and performed

sensitivity checks across alternative cut-off points to ensure the robustness of the findings. The specific values are shown in Table 7.

Table 7. Multiple Regression Predicting Academic Entitlement

Predictor	B	SE	β	t	p	95% CI
Constant	3.30	0.40	—	8.27	< .001	[2.52, 4.08]
Gender	-0.10	0.05	-0.09	-2.26	0.02	[-0.19, -0.02]
Age	-0.02	0.02	-0.05	-1.08	0.28	[-0.06, 0.02]
Grade	-0.02	0.03	-0.03	-0.64	0.53	[-0.07, 0.04]
Number of children	0.03	0.07	0.01	0.38	0.71	[-0.11, 0.16]
Discipline	-0.11	0.05	-0.09	-2.37	0.02	[-0.20, -0.02]

The results showed that the overall model was statistically significant, $F(5, 880) = 2.70$, $p = .02$, although the explained variance was relatively small ($R^2 = .02$, adjusted $R^2 = .01$). Among the predictors, gender ($\beta = -0.09$, $p = .024$) and academic discipline ($\beta = -0.09$, $p = .02$) were statistically significant. In contrast, age, grade, and number of children were not significant predictors ($ps > .05$). Overall, demographic variables demonstrated limited explanatory power for academic entitlement.

The findings of this study differ in notable ways from those of research conducted on non-mainland Chinese student populations, which represents an intriguing outcome. Greenberger et al. (2008) found that Asian students showed significantly higher levels of AE than their Caucasian counterparts. Similarly, Gwinner et al. (2021) reported that AE was both prevalent and relatively high among business students in Taiwan. Other cross-ethnic studies have also consistently identified elevated AE among Asian student samples. For instance, two separate studies by Ellis et al. (2021) revealed that Asian students displayed a small but statistically significant increase in AE compared to White students. In contrast, the present study, drawing on data from undergraduate students in Xinxiang City, Henan Province, found that only about one-third of participants demonstrated academic entitlement, with an overall low mean score. These findings suggest that students from this mainland Chinese sample may differ from Asian student populations examined in Western contexts. Cultural background may represent an important contextual factor underlying these differences.

4.2. Differential Analysis of Academic Entitlement in Demographic Variables

This study examined the effects of five demographic variables on academic entitlement. The results are presented in Table 8.

Table 8. Differences in Academic Entitlement Across Demographic Variables

	Demographic Variables		M	SD	t/F(df)	p	Effect size
AE	Gender	male	2.72	0.69	$t(884)=1.15$	0.25	$d=0.08$
		female	2.67	0.65			
	Grade	Freshman	2.69	0.68	$F(3,882)=4.36$	0.005	$\eta^2 = 0.02$
		Sophomore	2.75	0.65			
		Junior	2.78	0.65			
		Senior	2.50	0.67			
	Number of children	Only children	2.67	0.62	$t(884)=-0.38$	0.71	$d=-0.04$

		Non-only children	2.69	0.68			
	Birth order	First	2.67	0.65	F(3,882)=1.35	0.26	$\eta^2 = 0.005$
		Second	2.71	0.71			
		Third	2.77	0.59			
		Fourth	2.41	0.69			
	Academic discipline	Liberal Arts	2.74	0.65	t(884)=1.63	0.14	d=0.11
		Sciences	2.66	0.68			

Note. Independent-samples t tests were conducted for dichotomous variables, and one-way ANOVAs were conducted for variables with more than two categories. Effect sizes are reported as Cohen's d for t tests and partial eta squared (η^2) for ANOVAs.

First, this study found that gender had no significant effect on academic entitlement. In contrast, studies from other countries have yielded inconsistent results: some indicated that male students exhibited higher levels of AE than females (Greenberger et al., 2008; Ciani et al., 2008; Hartman, 2012), while others found that female students scored significantly higher than male students on AE (Al-Adamat et al., 2024). In any case, the research results indicate that there are significant differences between male and female college students. However, the results of this study indicate that male students scored slightly higher than female students among Chinese university students, but statistical tests revealed that this difference was not significant. Several other studies focusing on Chinese university students also reported no significant gender differences in AE (Lu & Jian, 2021; Shafait & Sahibzada, 2024).

According to the 2022 National Education Development Statistics Bulletin released by the Ministry of Education of the People's Republic of China (2023), among the students of regular undergraduate programs, the proportion of female students is 52.4%; among those of regular vocational colleges (dual degree programs), the proportion of female students is 50.6%. Overall for both undergraduate and vocational education, female students now make up the majority. Liang and Zhou (2023) reported that female students outperform their male counterparts in both academic and non-academic domains. This advantage is consistently observed across different majors and grade levels, and is particularly pronounced among female students in humanities disciplines and those enrolled in non-elite universities. Historically, Chinese traditional culture has valued sons over daughters. However, with the development of society, Chinese female college students have not only shown advantages in the participation in higher education, but also in academic achievements. Therefore, it is understandable that Chinese female college students' AE scores are lower than male students', but the difference is not significant.

Second, significant differences in AE were observed across grade levels. Prior to the analysis, Levene's test indicated that the assumption of homogeneity of variance was met, $F(3, 882) = 0.15$, $p = 0.93$. A one-way ANOVA revealed a significant main effect of grade on AE, $F(3, 882) = 4.36$, $p = 0.005$, $\eta^2 = 0.02$. Post hoc comparisons using the Tukey HSD test showed that senior (fourth-year) students reported significantly lower levels of AE than students in the other three grades, whereas no significant differences were found among freshmen, sophomores, and juniors. Existing literature presents two contrasting perspectives regarding the relationship between time and AE: one argues that older students possess a stronger sense of entitlement (Schaefer et al., 2013; Parker, 2017; Luckett et al., 2017), while the other suggests that younger students report higher AE (Huang, 2017). However, inconsistent findings—such as those from Lemke et al. (2017), who conducted two surveys at a liberal arts college and reported discrepant

results regarding gender differences in AE between 2009 and 2017—indicate that the effect of time on AE may be moderated by additional factors.

The lower level of academic entitlement among senior students may also be understood from a developmental perspective. As students approach graduation, they are increasingly exposed to external evaluation systems such as the job market and postgraduate entrance examinations. These experiences may be associated with a more realistic appraisal of their own abilities and reduced expectations of special academic treatment. In this sense, academic entitlement may reflect a dynamic and context-dependent construct rather than a stable individual trait.

Third, the number of children in the family had no significant effect on AE. Stiles et al. (2017) suggested that China's one-child policy made only children the center of their families and the focus of future expectations, leading parents to place higher academic demands on them. This pressure was thought to make only children more self-centered. Based on this, the present study hypothesized that only children in China would exhibit higher levels of AE than non-only children. However, the results showed no significant difference in AE between these two groups. One possible explanation is that in contemporary Chinese culture, children who attend university generally receive substantial family attention and support regardless of the number of siblings, which may diminish the effect of family size on AE. Another reason could be the relatively small sample size of only children ($n=88$) in this study, which may have limited the statistical power to detect existing differences.

Fourth, birth order had no significant impact on AE among university students. This variable was included based on Adler's theory of birth order, which proposes that individuals' personality and abilities differ by birth order, with first-borns often exhibiting greater leadership and competitiveness (Sultan & Malik, 2023). Wray-Lake et al. (2010) found that first-born children had more autonomy than their younger siblings. However, this study did not find birth order to be a significant factor influencing AE. It is possible that AE represents an attitude that stabilizes in early adulthood, rather than being shaped by university environment or short-term experiences.

Fifth, no significant differences in AE were found across academic disciplines. Jakiel (2020), in a study of a military academy, reported that students in science, health sciences, and professional colleges had the lowest AE scores, suggesting that stricter admission standards in these fields may filter out applicants with higher levels of AE. Additionally, Ellis et al. (2021) found that physician assistant students displayed lower AE compared to those in other majors. The discrepancy between these previous findings and the results of this study may be due to the broad categorization of disciplines in this study—only dividing into arts and sciences without further specialization—as well as the fact that the sample was drawn from comprehensive universities where management policies and student requirements are relatively uniform across disciplines, thereby reducing observable differences in AE.

4.3. Discussion

The present study examined the prevalence of academic entitlement and its relationship with demographic variables among Chinese university students. The findings indicate that although academic entitlement is not highly prevalent, a non-negligible proportion of students still exhibit moderate to elevated levels, suggesting the need for continued attention from educators and researchers. Importantly, demographic variables showed limited explanatory power in predicting academic entitlement, with one exception: senior students reported significantly

lower AE than their juniors, suggesting that developmental transitions (e.g., approaching graduation and facing job market pressures) may reduce entitled attitudes. This finding indicates that academic entitlement may be more strongly shaped by psychological and contextual factors rather than basic demographic characteristics. Such an interpretation aligns with previous research emphasizing the role of individual beliefs, motivational processes, and educational environments in shaping students' academic attitudes.

The small amount of explained variance further highlights the complexity of academic entitlement as a construct. It is likely influenced by a combination of cultural, familial, and individual psychological factors, rather than isolated demographic variables. In the Chinese cultural context, where collectivist values and respect for authority are traditionally emphasized, the emergence of academic entitlement may reflect ongoing social and educational transformations. The limited predictive power of demographic variables may also reflect the poor performance of the adapted AEQ in this sample. Several fit indices failed to meet acceptable standards, indicating weak model fit. This suggests that certain AEQ items may not be culturally appropriate for Mainland Chinese university students. Thus, the low academic entitlement scores observed in the Xinxiang sample may be partly due to poor item functioning rather than a true lack of entitlement. This underscores the need for culturally adapted measurement tools. These findings contribute to the literature by providing empirical evidence from a large Mainland Chinese sample, extending prior research that has predominantly focused on Western contexts. At the same time, the limited predictive power of demographic variables suggests that future research should focus more on psychological constructs such as academic self-efficacy, personality traits, and parenting styles.

5. Conclusion, Significance and Recommendation

5.1. Conclusion

This study, through an empirical investigation of the current situation of academic entitlement among Chinese college students, has reached the following main conclusions:

Firstly, this study found that the overall academic entitlement of Chinese college students is at a moderately low level, with the average score being lower than the theoretical median. More importantly, only about one-third of the students showed a certain degree of academic entitlement tendency. This finding suggests that academic entitlement may not be a widespread phenomenon among Chinese college students. This is in contrast to many Western studies that view academic entitlement as a common predicament in higher education. The discrepancy suggests that academic entitlement may not represent a universally shared belief across cultures and may be associated with specific cultural, educational, and social contexts.

Secondly, in the analysis of demographic variable differences, this study reveals the unique manifestation of academic entitlement within the Chinese cultural context. Apart from the grade variable, academic entitlement did not show significant differences in terms of gender, the number of children, birth order, and academic discipline. The uniqueness of this result lies in two aspects: Firstly, the absence of gender differences stands in contrast to some Western studies, suggesting that academic socialization in China may be associated with relatively smaller gender differences in entitlement attitudes. It challenges the stereotype repeatedly reported in Western literature that "male students have greater academic entitlement" (Greenberger et al., 2008). Similarly, the lack of effect of family structure variables

may be related to broader cultural emphases on educational investment across all children, rather than differential treatment based on birth order or only-child status. Secondly, the academic entitlement level of senior year students was significantly lower than that of students in other grades. This pattern may be associated with students' increasing awareness of their own abilities and responsibilities as they approach graduation and face the pressures of the job market or postgraduate entrance examinations, which coincides with lower levels of academic entitlement.

Third, these findings carry important implications for understanding AE within the Chinese cultural context. The relatively low prevalence and weak demographic differentiation of AE suggest that this construct should not be uncritically generalized from Western contexts to Chinese higher education. Instead, AE among Chinese students appears to be more situational and developmental rather than a stable or widespread attitudinal problem. From a practical perspective, these results caution against over-pathologizing Chinese college students' academic attitudes and suggest that educational interventions should be targeted toward specific developmental stages, particularly earlier years of university study. More broadly, the study underscores the importance of incorporating cultural context into both the measurement and interpretation of academic entitlement, thereby contributing to a more culturally sensitive understanding of this construct. Given the cross-sectional nature of the data, the findings should be interpreted as associations rather than causal relationships.

5.2. Significance

Based on the above conclusions, this study has both theoretical and practical significance.

5.2.1. Theoretical significance

This study contributes to the cross-cultural understanding of academic entitlement in two main ways:

1. It provides empirical evidence from an Eastern context, suggesting that academic entitlement may not be a universally prevalent phenomenon across cultures.
2. It highlights the importance of incorporating cultural values (e.g., collectivism, respect for teachers and elders) and educational systems (e.g., family and school education) into the theoretical framework of academic entitlement.
3. The study reveals potential cultural non-equivalence of a widely used AE scale (the AEQ) in the Chinese context. This finding calls for caution when directly applying Western-developed instruments to Eastern educational settings and underscores the need for culturally grounded measurement development.

5.2.2. Practical significance

The findings of this study offer several implications for higher education practitioners and administrators:

1. Strengthening academic responsibility education: Universities should continue to emphasize the cultivation of students' academic responsibility rather than treating academic entitlement as a widespread issue.

2. Enhancing awareness and early guidance: Educators should remain attentive to a small group of students who exhibit higher levels of academic entitlement and provide timely guidance.
3. Minimizing negative classroom impacts: Targeted strategies should be implemented to reduce the potential negative effects of academic entitlement on classroom dynamics, teacher–student relationships, and the overall educational environment.

5.3. Recommendation

This study has several limitations, which point to directions for future research.

First, this study adopted a cross-sectional design, which cannot confirm causal relationships behind the observed grade differences. Future research should employ longitudinal tracking designs to reveal the dynamic trajectory of students' academic entitlement from enrolment to graduation and its underlying causes.

Second, this study relied primarily on quantitative questionnaires, lacking in-depth description and analysis of the specific behavioral manifestations of academic entitlement among Chinese students (e.g., how they make requests to instructors). Subsequent research should introduce qualitative interviews or case studies to clarify the unique characteristics of academic entitlement among Chinese college students.

Third, although the sample size of this study was relatively large, all participants were drawn from a single, unnamed university in Xinxiang City, Henan Province—a mid-tier inland city in central China. This university is not among the nation's elite "985" or "211" institutions. Students in Xinxiang may differ substantially from those in coastal megacities (e.g., Shanghai, Guangzhou), elite universities, western interior provinces (e.g., Xinjiang, Gansu), or vocational education contexts. Therefore, the findings may not be directly generalizable to other regions or types of institutions. Future research should expand the sample to include more diverse regions and institutional types to test the replicability of these findings.

Fourth, future research should prioritize the cultural adaptation or revision of the AEQ for Mainland Chinese samples. The poor performance of specific items (e.g., AE2) suggests that some entitlement-related beliefs manifest differently in collectivist, authority-respecting educational cultures. Developing a context-appropriate scale—through qualitative exploration of how Chinese students express academic entitlement—would enable more accurate measurement and cross-cultural comparison.

Fifth, to better understand how academic entitlement manifests in Chinese students' daily academic lives, future research should conduct qualitative follow-up studies. For example, semi-structured interviews or focus groups could explore how students formulate requests to instructors, interpret course policies, and negotiate grades. Such qualitative data would complement the present quantitative findings by capturing the nuanced, context-dependent expressions of academic entitlement that standardized scales may miss. This would also inform the cultural adaptation of measurement tools.

In addition, although Henan Province has relatively low higher-education tuition levels compared to national averages (Henan Provincial People's Government, 2020) and comprehensive financial aid policies are in place, these contextual characteristics reflect only

a subset of China's diverse higher education landscape. Thus, the results are specific to university students in Xinxiang, Henan Province, and should not be generalized to students from other regions or institution types in China.

Acknowledgments

I am very grateful to Dr. Asyraf and Dr. Maizura for their many suggestions on my research direction and approach. I am also very grateful to the undergraduate students who participated in this research; their contributions made the work possible.

Statement on AI-Assisted Tools

AI-assisted tools were used exclusively for translation and language polishing of the authors' original Chinese manuscript. No artificial intelligence tools were involved in the study design, data analysis, or interpretation of the results. The authors take full responsibility for the content of this manuscript.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abdellatif, M. S. (2022). Moral intelligence and its relationship to academic entitlement and academic performance of secondary school students. *European Journal of Educational Research*, 11(4), 2291–2301. <https://doi.org/10.12973/eujer.11.4.2291>
- Al-Adamat, O., Husban, T., Tarawneh, H., Hemedan, N. S., & Shogran, R. I. (2024). Predictive ability of moral intelligence in academic entitlement among university students. *Educational Administration: Theory and Practice*, 30(5), 3295-3309. <https://doi.org/10.53555/kuey.v30i5.3440>
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.
- Blincoe, S., & Garris, C. P. (2017). Challenging the assumption of a western phenomenon: Academic entitlement in Saudi Arabia. *Journal of Experimental Education*, 85(2), 278–290. <https://doi.org/10.1080/00220973.2016.1199009>
- Boswell, S. S. (2012). "I deserve success": academic entitlement attitudes and their relationships with course self-efficacy, social networking, and demographic variables. *Social Psychology of Education*, 15(3), 353–365. <https://doi.org/10.1007/s11218-012-9184-4>
- Bronfenbrenner, U. (2000). Ecological systems theory. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology* (Vol. 3, pp. 129–133). Oxford University Press.
- Carollo, T. (2020). *Measuring the impact of academic entitlement in first-year college students*. <https://scholarworks.lib.csusb.edu/etd/1080>
- Chapin, T. M. (2018). *Associations between Academic Entitlement, Learning Approaches and Awareness of Future Consequences in Baccalaureate Nursing Students* [Doctoral dissertation, Capella University].
- Chowdhury, F. (2018). Grade inflation: Causes, consequences and cure. *Journal of Education and Learning*, 7(6), 86–92. <https://doi.org/10.5539/jel.v7n6p86>
- Ciani, K. D., Summers, J. J., & Easter, M. A. (2008). Gender differences in academic entitlement among college students. *Journal of Genetic Psychology*, 169(4), 332–344. <https://doi.org/10.3200/GNTP.169.4.332-344>
- Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement*, 12(4), 425-434.
- Ellis, J., Bacon-Baguley, T., & Otieno, S. (2021). Academic entitlement in physician assistant students. *The Journal of Physician Assistant Education*, 32(1), 1–9. <https://doi.org/10.1097/JPA.0000000000000335>

Finney, G. T., & Finney, Z. R. (2010). Are students their universities' customers? An exploratory study. *Education + Training*, 52(4), 276–291. <https://doi.org/10.1108/00400911011050954>

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1), 39-50.

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research*, 69(8), 3192-3198. <https://doi.org/10.1016/j.jbusres.2015.12.008>

Greenberger, E., Lessard, J., Chen, C., & Farruggia, S. (2008). Self-entitled college students: Contributions of personality, parenting and motivational factors. *Journal of Youth and Adolescence*, 37(10), 1193–1204. <https://doi.org/10.1007/s10964-008-9284-9>

Gruhlke, H. (2018). THE ETHICS OF GRADE INFLATION: INVESTIGATING THE ETHICAL IMPLICATIONS OF GRADE INFLATION IN US COLLEGES. *Business Journal for Entrepreneurs*, 2018(2).

Gwinner, T., Ludlum, M., Wiltz, A., & Carreno, J. (2021). IS ACADEMIC ENTITLEMENT A WESTERN CONCEPT? AN EXAMINATION OF TAIWANESE BUSINESS STUDENTS. *Global Journal of Business Pedagogy*, 5(1), 87–103. <https://doi.org/10.47177/gjbp.05.01.2021.087>

Hair, J. F., Black, W.C., Babin, B., Anderson, R, and Tatham, R. (2010). SEM: An introduction. *Multivariate data analysis: A Global Perspective*, 12(5), 629-686.

Halat, D. H., Al-Jumaili, A. A., Ahmed, K. K., Rahal, M., Hamad, A., Darwish, R. M., & Alsharif, N. Z. (2024). Academic entitlement among pharmacy students in the arab world: a multi-national exploratory study. *American Journal of Pharmaceutical Education*, 88(2), 100640. <https://doi.org/10.1016/j.ajpe.2023.100640>

Hartman, T. B. (2012). *An analysis of university student academic self-entitlement: Levels of entitlement, academic year, and gender* (Master's Theses, Eastern Michigan University). Digital Commons Network. <https://commons.emich.edu/theses/390>

Hassel, H., & Lourey, J. (2005). The death of student responsibility. *College Teaching*, 53(1), 2-13. <https://doi.org/10.3200/CTCH.53.1.2-13>

Henan Provincial People's Government. (2020, June 13). *Henan Province implements new tuition standards for new students in public universities this fall: New students who cannot afford tuition will be admitted first*. Henan Provincial Government Official Website. <https://www.henan.gov.cn/2020/06-13/1549095.html>

Hong, F. Y., Chiu, S. L., & Cheng, K. T. (2012). The study of academic entitlement: Development of measurement scale and the relationships between relational variables among Taiwan university students—Eight colleges in north and central Taiwan as an example. *Journal of Educational Research and Development*, 8, 141-174.

Hong, F. Y., Huang, D. H., Lin, M. P., & Lin, H. Y. (2016). Class social situation and cultural value prediction factors of the academic entitlement of college students. *Education and Urban Society*, 49(3), 341–360. <https://doi.org/10.1177/0013124516631623>

Hu, N. N., Lin, S. L., Zhao, D., He, Z., & Xu, P. (2021). The reliability and validity of the Chinese version of Academic Entitlement Questionnaire. *Chinese Journal of Nursing Education*, 18(12), 1110–1113.

Huang, S. (2017). *Psychological and academic entitlement: Psychosocial and cultural predictors and relationships with psychological wellbeing* [Doctoral dissertation, University of Windsor(Canada)]. Electronic Theses and Dissertations. <https://scholar.uwindsor.ca/etd/7266>

Huang, S., & Kuo, B. C. (2020). Demographic, psychosocial, and cultural predictors of entitlement in a multiethnic Canadian undergraduate sample. *Social Psychology of Education*, 23(2), 523–535. <https://doi.org/10.1007/s11218-020-09547-1>

Jakiel, A. J. (2020). *A Study of Self-reported Academic Entitlement Scores on a Senior Military College Campus* [Doctoral dissertation, Valdosta State University]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/openview/371ea3f26c4a739209093d0848d1d4c5/1?pq-origsite=gscholar&cbl=18750&diss=y>

Kinne, B. L., Goehring, M. T., & Williams, B. L. (2022). Academic entitlement and its potential educational consequences: A scoping review. *Journal of Physical Therapy Education*, 36(2), 115–121. <https://doi.org/10.1097/JTE.0000000000000231>

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Knepp, K. A., & Knepp, M. M. (2022). Academic entitlement decreases engagement in and out of the classroom and increases classroom incivility attitudes. *Social Psychology of Education*, 25(5), 1113–1134. <https://doi.org/10.1007/s11218-022-09716-4>

Kopp, J. P., Zinn, T. E., Finney, S. J., & Jurich, D. P. (2011). The Development and Evaluation of the Academic Entitlement Questionnaire. *Measurement and Evaluation in Counseling and Development*, 44(2), 105–129.

Kurtyilmaz, Y. (2019). Adaptation of academic entitlement questionnaire. *Anadolu Journal of Educational Sciences International*, 9(2), 314–351. <https://doi.org/10.18039/ajesi.577234>

Lemke, D., Marx, J., & Dundes, L. (2017). Challenging notions of academic entitlement and its rise among liberal arts college students. *Behavioral Sciences*, 7(4), 81. <https://doi.org/10.3390/bs7040081>

Liang, W., & Zhou, Y. (2023). Why Females Outperform Males? The Role of Non-cognitive Ability in Explaining the Gender Gap in College Students' School Performance. *In China Journal of Economics*.(01),344-374. <http://dx.doi.org/10.16513/j.cnki.cje.20230216.002>

Lin, M. H., & Pu, Y. S. (2018). Investigation and analysis of college students' understanding of cheating in exams. *Journal of Jilin Radio and Television University*, (5), 30.

Lippmann, S., Bulanda, R. E., & Wagenaar, T. C. (2009). Student entitlement: Issues and strategies for confronting entitlement in the classroom and beyond. *College Teaching*, 57(4), 197–204. <http://dx.doi.org/10.1080/87567550903218596>

Lu, M. P., & Jian, X. Z. (2021). The performance characteristics and influencing factors of the "academic entitlement" of Chinese college students. *Education Research Monthly*, (5), 78–82. <https://doi.org/10.16477/j.cnki.issn1674-2311.2021.05.012>

Luckett, M., Trocchia, P. J., Noel, N. M., & Marlin, D. (2017). A typology of students based on academic entitlement. *Journal of Education for Business*, 92(2), 96–102. <https://doi.org/10.1080/08832323.2017.1281214>

Lu, M. P., Li, X. Y. & Yang, Q. (2023). Analysis of strategies to improve the learning quality of college students in my country from the perspective of “academic privilege”. *Taste Classics*, (15), 132–134.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>

Mellor, J. K. (2011). *Academic entitlement and incivility: Differences in faculty and students' perceptions* [Doctoral dissertation, The University of Arizona]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/openview/650868f38b528b138c3436aed80eff9c/1?pq-origsite=gscholar&cbl=18750>

Ministry of Education of the People's Republic of China. (2023, July 5). *2022 National Education Development Statistics Bulletin*. https://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/202307/t20230705_1067278.html

Parker, L. L. (2017). *Competition and Academic Entitlement USA* [Doctoral dissertation, Walden University]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/openview/1652fdd45e421e8e6b4cf936774d406b/1?pq-origsite=gscholar&cbl=18750>

Peng, B., Pang, B., Lv, M., Li, H., Wang, X., & Hu, N. (2022). *The Chain Mediating Effects of External Locus of Control and Psychological Entitlement between Parental Over-Protection and Academic Entitlement among Nursing Students: A cross-sectional Study*. <https://doi.org/10.21203/rs.3.rs-2343227/v1>

Peng, B., Hu, N., Li, H., Pang, B., Lv, M., Wang, X., & Li, Y. (2024). A serial mediation model reveals the association between parental over-protection and academic entitlement among nursing students. *Scientific Reports*, 14(1), 15703. <https://doi.org/10.1038/s41598-024-66207-6>

Rao, T. K., & Kay, C. S. (2026). Ego and academic ethics: A brief investigation into the associations among three facets of narcissism, academic entitlement, and academic dishonesty. *Personality and Individual Differences*, 253, 113627. <https://doi.org/10.1016/j.paid.2025.113627>

Sang, X.C., Zhu, Y., Wu, J.Y., Ding, R., J., Lu, S.Y., Wen, S., & Ai, Z. (2023). Analysis of the status quo and influencing factors of academic entitlement among undergraduate nursing students. *Medical Higher Vocational Education and Modern Nursing*, 6(2), 93–97. <https://doi.org/10.3969/j.issn.2096-501X.2023.02.001>

Schaefer, T., Barta, M., Whitley, W., & Stogsdill, M. (2013). The "You owe me!" Mentality: A Student Entitlement Perception Paradox. *Journal of Learning in Higher Education*, 9(1), 79–91.

Shafait, Z., & Sahibzada, U. F. (2024). "I deserve better grades." Compliance-gaining perspective of dark triad traits, power distance and academic entitlement in Chinese higher education. *Kybernetes. Advance online publication*. <https://doi.org/10.1108/K-08-2023-1454>

Stiles, B. L., Pan, M., LaBeff, E. E., & Wong, N. (2017). The role of academic entitlement in college cheating: A comparison between China and the United States. *Research in Higher Education Journal*, 33, 1-15.

Stiles, B. L., Wong, N. C. W., & LaBef, E. E. (2018). College cheating thirty years later: The role of academic entitlement. *Deviant Behavior*, 39(7), 823–834. <https://doi.org/10.1080/01639625.2017.1335520>

Sultan, H., & Malik, S. (2023). Development of Adler Birth Order Scale. *Pakistan Journal of Psychological Research*, 38(2). <https://doi.org/10.33824/PJPR.2023.38.2.12>

Vallade, J. I., & Martin, M. M. (2014). Academic entitlement, grade orientation, and classroom justice as predictors of instructional beliefs and learning outcomes. *Communication Quarterly*, 62(5), 497–517. <https://doi.org/10.1080/01463373.2014.949386>

Wray-Lake, L., Crouter, A. C., & McHale, S. M. (2010). Developmental patterns in decision-making autonomy across middle childhood and adolescence: European American parents' perspectives. *Child Development*, 81(2), 636–651. <https://doi.org/10.1111/j.1467-8624.2009.01420.x>

Wu, D.G. (2019). *Accelerating the "Double First-Class" Construction and Realizing Connotative Development—Proceedings of the "2018 Higher Education International Forum Annual Conference"*. Lanzhou University Press.

Yu, G. L., & Li, S. (2020). The Concept of Academic Entitlement and Its Influence on Mental Health. *Renmin University of China Education Journal*, (1), 133–145.

Zare, M. (2022). The formation and management of academic entitlement: The case of extra credit assignments. *College Teaching*, 70(4), 405–412. <https://doi.org/10.1080/87567555.2021.1971601>

Zhou, S.S. (2023). *Analysis of the current situation and the pathway of influencing factors of academic entitlement among college and undergraduate nursing students* [Master's thesis, Xinxiang Medical University]. China Doctoral /Master's Theses Database. <https://doi.org/10.27434/d.cnki.gxxyc.2023.000253>

Appendix 1.**Items of the Academic Entitlement Questionnaire**

Item	English Version	Chinese Version
1	If I don't do well on a test, the professor should make tests easier or curve the grades.	□□□□□□□□□□□□□□□□ □□□□□□□
2	Professors should only lecture on material covered in the textbook and assigned readings.	□□□□□□□□□□□□□□□□
3	If I am struggling in a class, the professor should approach me and offer to help.	□□□□□□□□□□□□□□□□ □□□□□□□□
4	It is the professor's responsibility to make it easy for me to succeed.	□□□□□□□□□□□□□□□□
5	If I cannot learn the material for a class from lecture alone, then it is the professor's fault when I fail the test.	□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□
6	I am a product of my environment. Therefore, if I do poorly in class, it is not my fault.	□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□
7	I should be given the opportunity to make up a test, regardless of the reason for the absence.	□□□□□□□□□□□□□□□□□□
8	Because I pay tuition, I deserve passing grades.	□□□□□□□□□□□□□□□□□□

Note. The English version of the scale was developed by Kopp et al. (2011), and the Chinese version was adapted by Hu et al. (2021).

Appendix 2.**Comparison of CFA Models Before and After Removing AE2**

Model	χ^2	df	χ^2/df	CFI	TLI	IFI	NFI	SRMR	RMSEA	RMSEA 90% CI	ΔCFI	$\Delta RMSEA$
8-item model	138.62	20	6.93	.91	.88	.91	.900	—	.08	[.07, .10]	—	—
7-item model (AE2 removed)	135.12	14	9.65	.91	.87	.91	.902	.06	.10	[.08, .11]	-0.002	+0.017

A Descriptive Study of Assertiveness, Social Skills, and Psychological Well-Being Among First-Year Students in Malaysian Public Universities

ABSTRACT

Authors: Harizah Izyan Samsudin¹, Muhammad Asyraf Che Amat^{1*}, Salleh Amat² and Abu Yazid Abu Bakar²

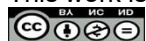
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

²Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

*Corresponding author:
mhdasyraf@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This study investigates the levels of assertiveness, social skills, and psychological well-being among first-year undergraduate students in Malaysian public universities (UA), with a focus on understanding their psychosocial readiness during the transition into higher education. A quantitative cross-sectional descriptive survey design was employed, involving 1,206 respondents selected through a multistage sampling procedure incorporating cluster and stratified sampling techniques across multiple public universities. Data were collected using three established instruments: the Rathus Assertiveness Schedule (RAS), the Social Skills Inventory (SSI-Del-Prette), and Ryff's Psychological Well-Being Scale, all of which demonstrated strong internal consistency (Cronbach's alpha = .93, .94, and .88, respectively). Descriptive analyses revealed that students reported moderate levels of assertiveness ($M = 79.87$, $SD = 18.67$), social skills ($M = 86.49$, $SD = 24.49$), and psychological well-being ($M = 136.13$, $SD = 38.38$) across measured dimensions. These findings suggest that while students possess foundational psychosocial competencies, there remains substantial room for further development, particularly in adapting to the academic, social, and emotional demands of university life. The results highlight the transitional nature of the first-year experience, where students are still adjusting to new environments, expectations, and interpersonal contexts. This study contributes baseline empirical evidence on the psychosocial profiles of Malaysian first-year undergraduates and underscores the importance of structured institutional support. The findings may inform the design of targeted interventions, such as orientation programmes, peer mentoring, and skills-based training, aimed at strengthening students' psychosocial readiness and overall well-being in higher education settings.

Keywords: Assertiveness, Social Skills, Psychological Well-Being, University Students, Cross-Sectional Survey, Malaysia, Soft Skills

1. Introduction

Universities play a central role in society by disseminating knowledge, producing professionals with scientific expertise, and providing technical and intellectual support for cultural, social, and economic development (Ministry of Education, 2024). Beyond their academic mission, universities are also entrusted with nurturing students' holistic development. This responsibility aligns with the National Philosophy of Education, which emphasizes balanced growth across intellectual, spiritual, emotional, and physical domains to develop graduates who can contribute meaningfully to personal, societal, national, and religious advancement (Ministry of Education Malaysia, 2008; FPK, 2013). As a result, higher education should not only emphasize academic excellence but also promote student satisfaction and psychological well-being. Despite this perspective, research also consistently demonstrates that undergraduate students experience higher levels of stress, anxiety, and depression than the general population, a trend observed globally and within Malaysia (Cho et al., 2021; Labrague and De los Santos, 2020). Local studies also reported moderate to severe levels of depression, anxiety, and stress among Malaysian undergraduates (Ranita et al., 2019), while more recent findings indicate that students' psychological wellbeing remains at a moderate level (Aziz et al., 2024). These mental health challenges have been shown to negatively affect academic performance (Mahdavi et al., 2023), social relationships (Freitas et al., 2022), and overall quality of life (Slimmen et al., 2022).

Such concerns are particularly pronounced among first year undergraduates, who face the dual challenge of adapting to increased academic demands while navigating new social and cultural environments (Mohzana, 2024). During this key changeover period, students are expected to show more independence, initiative, and emotional maturity. In addition to academic competency, effective social skills are required for managing peer relationships and meeting social expectations (Blegur, Haq, & Barida, 2023; Toktas, Demir, & Barut, 2022; Suleman et al., 2022; Vanhove, Opdecam, & Haerens, 2024; World Economic Forum, 2025). However, research indicates that many Malaysian undergraduates are passive in academic and social settings, frequently avoiding engagement, withholding opinions, and lacking confidence in expressing or defending their ideas (Ang, 2016; Kipli et al., 2022; Sieng and Yussof, 2017). Assertiveness, as a key dimension of social skills, enables individuals to communicate their needs and viewpoints clearly while respecting the rights of others. It has been associated with greater resilience against bullying, constructive conflict resolution, and improved academic outcomes (Alberti and Emmons, 2017; Blegur et al., 2023).

Nevertheless, empirical evidence indicates that Malaysian undergraduates generally demonstrate low to moderate levels of assertiveness, which may constrain their confidence, critical thinking, and engagement compared with international peers (Ang, 2016; Kipli et al., 2022; Sieng and Yussof, 2017). Limited assertiveness may also hinder the development of healthy interpersonal relationships, thereby increasing vulnerability to stress and social isolation (Moss et al., 2021). Psychological wellbeing is equally central to students' adjustment and success in higher education. Ryff conceptualized psychological wellbeing as comprising autonomy, environmental mastery, personal growth, purpose in life, positive relations, and self-acceptance (Ryff, 1989, 2018). Higher psychological well-being has been linked to greater resilience, adaptability, and academic achievement (Korda et al., 2025). However, national surveys continue to report concerning levels of stress, anxiety, and depression among young people in Malaysia (Institute for Public Health, 2015, 2023). Despite the expanding focus on student mental health in higher education, existing research in Malaysia has predominantly examined psychological distress indicators such as stress, anxiety, and depression in isolation, with limited integration of positive psychosocial competencies. While studies have documented moderate to severe mental health challenges among undergraduates, far fewer have explored the psychosocial skills that enable students to cope effectively with academic and social demands.

In particular, assertiveness and social skills, which are closely associated with psychological wellbeing and adaptive functioning, remain under examined within the Malaysian higher education context. Moreover, much of the existing literature also focuses on undergraduate populations in general, thereby overlooking the unique vulnerabilities of first year students. The transition into university represents a

critical developmental period characterized by increased academic expectations, new social roles, and greater personal responsibility. First year undergraduates are therefore at heightened risk of adjustment difficulties, yet empirical evidence addressing their psychosocial readiness, particularly within Malaysian public universities, remains limited. Many existing studies rely on single institution samples or focus on narrow constructs, thereby constraining the generalizability and practical relevance of their findings. Understanding the levels of these psychosocial factors is essential for informing the development of appropriate support strategies for first-year students. However, empirical evidence that provides a comprehensive description of assertiveness, social skills, and psychological well-being within a multi-institution Malaysian context remains limited. Without such baseline data, higher education institutions may lack sufficient information to design targeted and contextually appropriate initiatives to support students' holistic development.

Therefore, the present study examines the levels of assertiveness, social skills, and psychological well-being among first-year undergraduates in Malaysian public universities. By providing a descriptive overview of these key psychosocial constructs, this study offers empirical insights into students' psychosocial readiness during the transition to university life.

Importantly, this study extends existing research by drawing on a large dataset ($N = 1,206$) collected from first-year students across five Malaysian public universities representing multiple geographical zones. In addition, the study incorporates the simultaneous assessment of assertiveness, social skills, and psychological well-being using three established instruments, namely the Rathus Assertiveness Schedule (RAS), the Social Skills Inventory (SSI-Del-Prette), and Ryff's Psychological Well-Being Scale. Collectively, these contributions provide baseline evidence that may support future research and inform the development of early and preventative support initiatives within higher education contexts.

2. Theoretical Framework

This study is grounded in established theoretical perspectives that support the inclusion of assertiveness, social skills, and psychological well-being as key components of students' psychosocial development in higher education contexts. Although the present study adopts a descriptive design and does not aim to test causal or predictive relationships among variables, the integration of relevant theoretical perspectives is essential to provide conceptual coherence and justify the selection of these constructs. This study is informed by Albert Bandura's Social Cognitive Theory (1977), which emphasises the interaction between personal beliefs, behavioural capabilities, and social environments in shaping human functioning. Within this perspective, assertiveness reflects individuals' confidence in expressing thoughts and needs in social contexts, aligning with the concept of self-efficacy. Social skills, in turn, represent learned behavioural competencies that facilitate effective interpersonal interaction and adaptation, particularly during transitional phases such as entry into university.

In addition, this study draws on Ryff's Model of Psychological Well-Being (1989), which conceptualises psychological well-being as a multidimensional construct comprising autonomy, environmental mastery, personal growth, purpose in life, positive relations with others, and self-acceptance. The dimension of positive relations with others highlights the importance of interpersonal functioning, suggesting that social competencies such as assertiveness and social skills are conceptually relevant to individuals' overall psychological well-being. Collectively, these theoretical perspectives provide a coherent basis for understanding assertiveness, social skills, and psychological well-being as interrelated dimensions of students' psychosocial functioning. While the present study does not empirically examine the relationships among these variables, the framework supports their inclusion as complementary constructs that reflect students' psychosocial readiness during the transition to higher education.

3. Literature Review

3.1. Assertiveness

Assertiveness is widely recognised as a fundamental communication and psychosocial skill that enables individuals to express thoughts, emotions, and needs clearly while maintaining respect for the rights of others. Effective communication plays a central role in sustaining healthy interpersonal relationships and reducing conflict, particularly within academic and social contexts (Rosenfeld et al., 2000). Conceptually, assertiveness is positioned on a behavioural continuum between passive and aggressive communication styles, representing a balanced mode of interaction that promotes self-expression without hostility or submission (Lazarus & Folkman, 1984). Rather than being a fixed personality trait, assertiveness is increasingly understood as a learnable social competence that can be developed through experience and structured intervention.

Early conceptualisations described assertiveness as an individual disposition (Salter, 1949), but subsequent theoretical developments reframed it as a core social skill associated with psychological adjustment and adaptive functioning. Lazarus (1971) characterised assertive behaviour as an indicator of social competence, while non-assertiveness was conceptualised as a social skills deficit. Further elaboration identified key assertive abilities, including the capacity to communicate personal needs openly, refuse unreasonable demands, express both positive and negative emotions appropriately, and manage interpersonal interactions effectively (Lazarus, 1973). These principles were later operationalised by Rathus (1973), who proposed behavioural dimensions of assertiveness that emphasise mutual respect, emotional expression, the right to make mistakes, the right to say “no,” and open communication. This framework has been widely applied in educational and psychological research due to its relevance across diverse social contexts.

However, empirical evidence from Malaysia suggests that university students generally exhibit low to moderate levels of assertiveness, despite its inclusion as a key component of generic or soft skills within higher education policy frameworks (Ang, 2016; Kipli et al., 2022; Ministry of Higher Education Malaysia, 2024). These findings indicate that while students may possess cognitive and academic capabilities, deficiencies in assertiveness may constrain their ability to engage effectively in academic and social environments. Similar patterns have been observed among professional students, including medical undergraduates, where moderate assertiveness levels were found to have implications for emotional regulation, decision-making, and professional communication (Alyssa Sureyya et al., 2021).

Assertiveness has also been identified as a key factor in students’ adjustment to university life. Research examining student adaptation indicates that assertiveness is more strongly associated with social and emotional adjustment than with academic performance alone, suggesting its primary influence lies in interpersonal functioning and psychological resilience (Rahayu, 2023). Additionally, low assertiveness has been linked to increased reliance on digital avoidance behaviours, such as excessive smartphone use, which may further reduce face-to-face social engagement and reinforce social withdrawal (Parmaksiz, 2019). These findings suggest that assertiveness plays a critical role in students’ social and emotional adjustment, highlighting its potential link with broader psychosocial competencies such as social skills and psychological well-being.

Despite the extensive international literature, research on assertiveness in Malaysian higher education remains limited. This gap restricts the ability of institutions to implement evidence-based strategies that support students’ psychosocial development and psychological well-being. Accordingly, this research is warranted to systematically assess assertiveness levels among Malaysian undergraduates. Addressing this gap may contribute to improved student well-being, academic engagement, and graduate employability, in line with national higher education objectives. These findings suggest that assertiveness plays a critical role in students’ social and emotional adjustment, highlighting its potential link with broader psychosocial competencies such as social skills and psychological well-being.

3.2. Social Skills

Social skills have evolved from early behavioural and social competence theories. Salter (1949) initially framed social competence as a behavioural tendency, while later work by Zigler and Phillips (1961) emphasised its role in adaptive functioning and social adjustment. Subsequent theoretical developments conceptualised social skills as learned, goal-directed behaviours that enable individuals to interact effectively and maintain positive interpersonal relationships (McFall, 1982). Contemporary perspectives highlight that social skills comprise both observable behaviours and underlying cognitive and emotional processes. These skills include verbal and non-verbal communication, emotional regulation, empathy, assertiveness, cooperation, and self-control, all of which operate within specific social and cultural contexts (Del Prette & Del Prette, 2017). Social skills are therefore not static traits but dynamic competencies that can be acquired and enhanced through learning, practice, and social interaction across the lifespan.

Within higher education settings, social skills play a critical role in students' academic engagement, psychological well-being, and social adjustment. Students with higher levels of social competence are more likely to participate actively in classroom discussions, collaborate effectively in group work, and communicate confidently with peers and lecturers (Toktas, Demir, & Barut, 2022). In contrast, limited social skills have been associated with social anxiety, avoidance behaviours, reduced academic involvement, and poorer mental health outcomes (Rahayu, 2023). This indicates that social skills are closely related to students' ability to function effectively in both academic and social environments, and may contribute to their overall psychological well-being.

The significance of social skills extends beyond academic contexts to labour market outcomes. Graduate unemployment in Malaysia has increasingly been attributed to deficiencies in soft skills rather than a lack of academic qualifications (Ministry of Higher Education Malaysia, 2024). Employers also consistently prioritise communication, teamwork, adaptability, and emotional intelligence over technical expertise alone (World Economic Forum, 2025). However, higher education institutions continue to emphasise academic achievement, often providing limited structured opportunities for systematic social skills development. Study conducted in Malaysian universities indicate that undergraduates generally demonstrate moderate levels of social skills, suggesting the presence of basic competencies but limited application in more complex interpersonal situations (Chin, 2016). These findings underscore the central role of social skills in supporting both individual functioning and successful social integration.

3.3. Psychological Well-Being

Psychological well-being (PWB) is widely conceptualised as a multidimensional construct encompassing individuals' optimal psychological functioning and positive engagement with life. Seminal work by Ryff conceptualises PWB through six core dimensions: self-acceptance, positive relations with others, autonomy, purpose in life, personal growth, and environmental mastery (Ryff & Singer, 2006; Ryff, 2018). This eudaimonic perspective emphasises meaning, self-realisation, and adaptive functioning rather than transient emotional states. Contemporary scholars further describe psychological well-being as the integration of positive affect, effective functioning, and meaningful life engagement (Ruggeri et al., 2020), bridging both hedonic and eudaimonic traditions (Tang et al., 2019).

Psychological well-being is not limited to emotional satisfaction but reflects individuals' capacity to manage life challenges, maintain healthy relationships, and pursue personal growth in a balanced manner. Individuals with high levels of PWB tend to perceive their lives as meaningful, display optimism, possess a positive self-concept, and demonstrate resilience in the face of stressors (Tobin & Neugarten, 1961; Suvera, 2016). Empirical evidence further indicates that psychological well-being is associated with favourable physical health outcomes, including lower risk of chronic illness and

increased longevity (Hernandez et al., 2018). Collectively, these findings suggest that PWB represents a holistic indicator of mental health that extends beyond the absence of psychological distress.

Within higher education contexts, psychological well-being has emerged as a critical concern, given increasing reports of mental health challenges among university students globally and in Malaysia. Although earlier local studies predominantly measured mental health through distress-based indicators such as depression, anxiety, and stress using instruments like the DASS-21, these approaches provide only a partial understanding of students' psychological functioning (Shahira et al., 2018; Manap et al., 2019). Findings from these studies reveal alarmingly high levels of anxiety and moderate levels of depression and stress among Malaysian undergraduates, particularly among female students and those in their final year of study, highlighting the psychological vulnerability of this population.

More recent research has adopted a positive psychology framework by directly assessing psychological well-being using multidimensional models such as Ryff's Psychological Well-Being Scale. Studies conducted among Malaysian university students report that most students demonstrate moderate levels of PWB, with significant variations across dimensions such as autonomy, purpose in life, and environmental mastery (Kunjiapu & Kunasegaran, 2021; Aziz et al., 2024). These findings suggest that while students may function adequately, many struggle with long-term purpose, emotional regulation, and life direction, particularly under academic and career-related pressures.

Systematic reviews further reinforce concerns regarding student psychological well-being in Southeast Asia. Dessauvage et al. (2022), in a PRISMA-guided systematic review, reported high prevalence rates of depression and anxiety among university students across the region, alongside limited evidence-based institutional interventions. The review highlights structural challenges such as stigma, insufficient mental health professionals, and fragmented psychosocial support systems, underscoring the need for preventive and developmental approaches rather than crisis-driven responses. The transition into university life, particularly during the first year of study, has been identified as a critical period for psychological well-being. Conceptual analyses indicate that first-year students frequently experience academic overload, social isolation, financial strain, identity confusion, and reluctance to seek help due to stigma (Oramas, 2024). Without early and systematic support, these stressors may undermine long-term psychological functioning and academic persistence. Consequently, scholars increasingly advocate for early, institution-wide interventions that integrate emotional support, peer mentoring, and mental health literacy into the university ecosystem.

Beyond individual vulnerabilities, psychological well-being is shaped by multiple contextual and psychosocial factors. Empirical studies demonstrate that social support from family, peers, and institutions, positive learning environments, and academic engagement significantly predict students' psychological well-being (Chaudhry et al., 2024). Similarly, psychological well-being has been consistently linked to academic outcomes, with students exhibiting higher well-being demonstrating greater self-regulation, motivation, and academic achievement (Xiang et al., 2024). Conversely, psychological distress negatively impacts academic performance, although achievement motivation may buffer these adverse effects (Mahdavi et al., 2023). Maladaptive psychological traits such as academic procrastination and maladaptive perfectionism have also been identified as significant predictors of reduced psychological well-being. Research indicates that excessive fear of failure and unrealistic self-expectations contribute to procrastination behaviours, which in turn exacerbate stress, guilt, and emotional exhaustion among students (Ahmad & Munir, 2022). Overall, the reviewed literature indicates that Malaysian university students generally exhibit low to moderate levels of well-being, with significant implications for mental health, academic success, and long-term development. These findings highlight the importance of addressing psychological well-being patterns alongside emotional well-being in student support initiatives.

3.4. Summary of Literature

Collectively, the reviewed literature indicates a consistent pattern in which Malaysian university students demonstrate moderate levels of assertiveness, social skills, and psychological well-being. While these constructs have been examined separately in prior studies, they represent interconnected dimensions of students' psychosocial adjustment, particularly during the transition to university life. Limited assertiveness may restrict students' ability to express their needs and participate actively in academic and social environments, while moderate levels of social skills may reduce opportunities for effective communication, collaboration, and peer support. These constraints may, in turn, contribute to only moderate levels of psychological well-being, especially in areas related to emotional regulation, purpose in life, and interpersonal relationships.

Furthermore, these patterns are influenced by broader contextual factors, including academic pressures, cultural norms that may discourage open self-expression, and limited structured opportunities for psychosocial skill development within higher education institutions. As a result, students may possess basic competencies but face challenges in applying them effectively in complex real-life situations. Therefore, a comprehensive understanding of these psychosocial competencies is necessary to provide baseline evidence for early intervention and support strategies. This justifies the present study's focus on examining the levels of assertiveness, social skills, and psychological well-being among first-year undergraduates in Malaysian public universities.

4. Methodology

4.1. Research Design

The survey employed a cross-sectional research design, with a structured questionnaire used as the primary data collection instrument. This design is commonly utilised in social science research as it allows for precise quantitative descriptions of population characteristics through the examination of a representative sample within a limited period (Sekaran, 2005). Cross-sectional surveys also enable the efficient collection of primary data prior to the application of inferential statistical analyses (Cohen et al., 2018). The survey was conducted to examine levels of assertiveness, social skills, and psychological well-being among first-year undergraduate students at Malaysian public universities (UA). First-year undergraduates were selected as the target population due to their transitional phase into higher education, a period often characterised by significant psychosocial adjustment.

4.2. Population

Population and sampling form the foundational, critically important elements of any research project. The appropriate population selection establishes the scope of the investigated issue and the scale of required data collection and analysis. Thus, researchers must carefully determine the population and sample as an initial step. Mohd Najib (2009) explains that proper population choice affects study effectiveness in problem definition and data acquisition. This research targets first-year undergraduate students from the 2021/2022 intake at Malaysian public universities. Selecting the first-year undergraduate students is significant, as they exhibit lower exposure to soft skills than those in subsequent years. They are also navigating the transition from school to university life, marked by social and psychological adjustments that impact self-adaptation, social abilities, assertiveness, and mental health. Astin (1993) and Pascarella and Terenzini (2005) underscore the first year as a vital period for soft skill development and psychosocial adjustment.

4.3. Sampling Procedure

Wiersma and Jurs (2009) describe a sample as a portion or subset of a larger group known as the population. Creswell (2014) notes that cluster random sampling can be applied across multiple

hierarchical levels of population clusters to meet researcher needs, such as students' fields of study at university. In this study, the sample comprises first-year undergraduate students who represent their academic year, assuming all have undergone adjustment to their respective programs. The employed sampling technique is stratified random sampling, combining cluster random sampling at the first stage and stratified random sampling at the second stage. Although various sampling methods exist, the researcher selected cluster random sampling for the survey data collection, following Cohen et al. (2017), who highlight its time-saving benefits and reduction of logistical issues by grouping the population. Stratified sampling, meanwhile, enables proportional division by specialization areas.

Additionally, cluster sampling was chosen due to the uneven distribution of first-year public university (UA) students across zones and the scattered placement of study samples across 20 UAs in 14 states on Peninsular Malaysia, Sabah, and Sarawak. This aligns with Creswell (2014), who recommends cluster sampling for large, geographically dispersed populations. Wiersma and Jurs (2008) further advice including all members of selected clusters in the sample. Therefore, drawing from Wiersma and Jurs (2009), the researcher applied cluster random sampling to select groups representing the population of first-year undergraduate students at Malaysian public universities.

For cluster random sampling, the first step the researcher took was to identify clusters that could represent the population before proceeding with the sampling process. Therefore, the researcher grouped the 20 UAs in Malaysia according to zones established by the Ministry of Higher Education. Based on the diagram, there are six UAs in the Central Zone, four UAs in the Southern Zone, four UAs in the Eastern Zone, four UAs in the Northern Zone, and two UAs in the Borneo Zone. After grouping the UAs by zone, the researcher conducted the selection of clusters by adapting the basic elementary sampling marble from jar technique (Neuman, 2009). This technique required the researcher to prepare a list of universities by zone. Then, the selection process was carried out by randomly choosing a university to represent each zone.

For this survey study, the researcher selected one university from each zone, considering that the Borneo Zone only had two UAs. As a result, five universities were selected using the elementary sampling marble from jar technique, namely Universiti Teknologi Mara (UiTM), Universiti Malaysia Pahang (UMP), Universiti Teknikal Malaysia Melaka (UTeM), Universiti Sains Malaysia (USM), and Universiti Malaysia Sabah (UMS). Choosing one university from each zone helps ensure that different geographic regions are represented in the sample. When applying this technique, the selection process was carried out only once. The selected universities represent different geographic zones and institutional categories, providing broad coverage of the Malaysian higher education context. However, as only one university was selected from each zone, the sample does not fully capture the variability across all public universities in Malaysia. Therefore, the findings should be interpreted with caution when generalised beyond the sampled institutions. Additionally, the tabulation of results based on the zones will give an overall picture that the sample for the study is widespread.

Before determining sample size, the researcher emailed and phoned all 20 Malaysian UAs to obtain estimated first-year undergraduate student populations for the 2021/2022 academic session. Based on the data obtained, the total number of first-year undergraduate students for the 2021/2022 academic session across the 20 UAs in Malaysia was 60,976. However, this study's sample draws from first-year undergraduates at just the five selected universities (UiTM, UMP, UTeM, USM, and UMS), totaling approximately 11,803. The students from these five universities provide representation across multiple geographic zones; however, the findings are limited to the sampled universities and should be interpreted with caution when generalised to all public universities in Malaysia.

4.3.1. Sample Size Determination

In determining the sample size, based on the Krejcie and Morgan (1970) for a population of 11,803 the

5% margin of error. Additionally, at least 10% was added to account for potential incomplete, damaged, or unreturned questionnaires (Mohd Najib, 2005). Therefore, the proposed sample size was 415 students. However, for this study, the researcher used proportional stratified sampling to determine the sample size required from these five universities. Stratified sampling is a method that requires all samples to be grouped according to certain parameters, and samples are selected from each group rather than randomly from the entire population.

According to Creswell (2005), the proportional stratified random sampling technique is suitable for populations that are not uniform. For this study, the sample was taken from the population of students at UiTM, UMP, UTeM, USM, and UMS, which have different backgrounds, locations, and university categories. USM is a research university, while UiTM and UMS are comprehensive universities. On the other hand, UTeM and UMP are focused universities. Hence, in this case, using simple random sampling would result in insufficient samples for certain categories. In proportional stratified sampling, the sample size determined is proportional to the groups formed. Since the division of groups strata and the number of samples taken to represent the entire population depends on the researcher, there is no specific formula for the stratified sampling process. However, the formula below has been widely used to calculate the sample size for each group (Cresswell, 2005):

$$\text{Sample Size} = \frac{\text{Total Sample}}{\text{Total Population of Group}} \times \text{Population of Sub Group}$$

For this study, the researcher decided to select 10% from each group. Therefore, the required sample size is 1,180 students, which includes 275 students from UiTM, 210 students from UTeM, 253 students from USM, 207 students from UMP, and 236 students from UMS, as discussed in Table 1 below. This sample size is sufficient to represent the population because based on Fraenkel and Wallen (1993), a larger sample size is better for representing a population.

Table 1. Sample Size by University

Zone	Public University	Population	Sample
Central	Universiti Teknologi Mara Shah Alam (UiTM)	2750	275
Southern	Universiti Teknikal Malaysia Melaka (UTeM)	2095	209
Northern	Universiti Sains Malaysia (USM)	2360	236
Eastern	Universiti Malaysia Pahang (UMP)	2068	207
Borneo	Universiti Malaysia Sabah (UMS)	2530	253
Total		11,803	1,180

4.3.2. Respondent Selection

Based on the total population of 15,583 first-year students across the five selected Malaysian public universities, the minimum required sample size was determined using the Krejcie and Morgan (1970) table. According to this guideline, a minimum of approximately 375 respondents is required for a population of this size. To enhance representativeness across institutions and support proportional allocation by university, the study targeted a larger sample. As a result, a total of 1,206 respondents were included in the final analysis, exceeding the minimum requirement. This larger sample strengthens the stability and representativeness of the descriptive findings.

The selection process was carried out systematically and transparently to ensure fair representation and reliable data. The first step involved creating an alphabetical list of first-year undergraduate students from the five selected universities. Systematic random sampling was then applied by selecting respondents at regular intervals (e.g., every 4th, 8th, 12th, and so on) until the required sample size was achieved for each university. By using this method, the study generated data that can be generalised to the sampled universities, with caution when extended beyond this context. This approach helps to

reduce selection bias and improve the reliability of the results (Flick, 2018). Before participating in the study, all participants were provided with detailed information about the study's purpose and procedures within the Google Form. They were informed that participation was voluntary and that they could withdraw at any time without consequence. By proceeding with the questionnaire, participants provided 1,226 responses were received. After data screening, 1,206 completed questionnaires met the necessary criteria, resulting in a usable response rate of approximately 90.7%. This larger sample size enhances the stability and reliability of the descriptive findings within the sampled universities.

4.3.3. Sampling Procedure Diagram

The detailed sampling procedure is illustrated in Figure 1 below. The diagram shows how the sampling process is carried out, from identifying the population to selecting the universities and finally selecting the respondents.

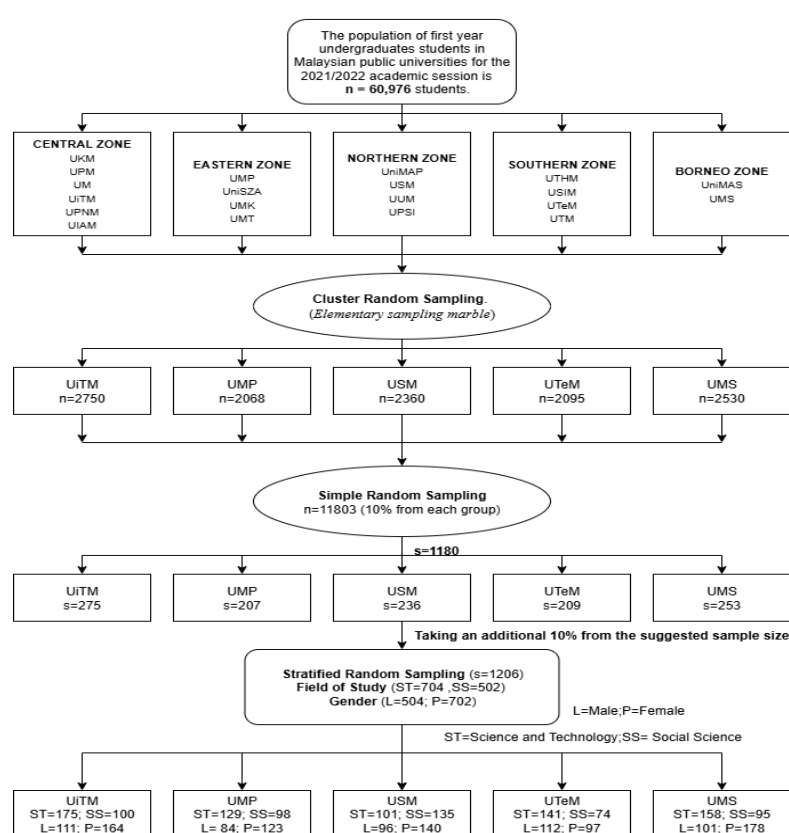


Figure 1. Sampling Procedure Diagram

4.4. Research Instrument

Research instruments are tools designed specifically for data collection purposes. To ensure comprehensive and convincing research, various instruments can be used (Fraenkel et al., 2018; Meriam & Tisdell, 2015). In this study, the researcher used a questionnaire as the instrument to collect the necessary data to answer the research questions and hypotheses. According to Fraenkel et al. (2012), a questionnaire can either be developed independently or adapted from existing instruments to save time. For this study, the questionnaire was developed using an "adapt and adopt" method. Specifically, the questionnaire includes four main sections which are Section A: Respondent profile, covering gender,

ethnicity, field of study, residence, and socioeconomic status. Section B focuses on assertiveness, using the Rathus Assertiveness Schedule (RAS). Section C addresses social skills, adapted from the Social Skill Inventory (Del-Prette, 2001). Section D assesses psychological well-being, based on Ryff's Psychological Well-Being Scale. The details of the research instrument are discussed in Table 2 below.

Table 2. Research Instrument

Section	Instrument / Item	Number of Items	Reference
A	Respondent Profile	6	-
B	Rathus Assertiveness Schedule (RAS)	30	Rathus (1973,1978); Hawa (2014); Abdullah (2005)
C	Social Skill Inventory Del-Prette (SSI)	38	Del Prette & Del Prette (2001;2013); Nor Azima (2008)
D	Ryff Psychological Well-Being Scale	42	Ryff (1989,2019); Hasimah (2010); Salina & Rahimi (2013); Norsayyidatina (2017)

4.4.1. Rathus Assertiveness Schedule (RAS)

Assertiveness in this study was measured using the Rathus Assertiveness Schedule (RAS), developed by Spencer Rathus (1973; 1978). The instrument, developed by Rathus (1973) was designed to assess levels of assertiveness among college students and adults. In this study, the researcher employed the Malay version of the RAS, which was adapted from previous studies conducted by Abdullah (2005), Zainol Abidin (2016), Hawa (2014), and Bakri (2016). Abdullah (2005) used the RAS to measure assertive behaviour among employees of the Johor District Customs Department. Meanwhile, Hawa (2014) utilised the RAS to examine the relationship between assertive behaviour and organisational commitment among support staff at Universiti Tun Hussein Onn Malaysia (UTHM).

The RAS instrument used in this study consists of 30 items representing five dimensions of assertive behaviour which is the right to mutual respect, the right to mutual feelings, the right to make mistakes, the right to say "no," and the right to ask questions and communicate. Fourteen items measure assertive behaviour, while sixteen items assess non-assertive behaviour. Responses are rated on a five-point Likert scale: 1 = very untrue; 2 = untrue; 3 = partly true and partly untrue; 4 = true; and 5 = very true. The total assertiveness score ranges from 30 to 150 for the 30 items. The minimum possible score is 30 (1×30), and the maximum possible score is 150 (5×30). The level of assertiveness, according to Rathus' (1976) classification, is divided into several categories based on the score obtained, namely aggressive, assertive, moderately assertive, less assertive, and passive (Rathus, 1976). For this study, the RAS, adapted from Abdullah' study (2005), categorizes assertiveness based on the score obtained, with the classifications being aggressive (121-150), assertive (91-120), moderately assertive (60-90), less assertive (30-59), and passive (0-29) as shown in the table 3 below. This classification is used to assess the extent to which an individual demonstrates assertiveness in social interactions and self-communication (Rathus, 1976).

Table 3. Interpretation of Assertiveness Score

Score	Level
0-29	Passive
30-59	Less Assertive
60-90	Moderately Assertive
91-120	Assertive
121-150	Aggressive

Next, to determine the level of assertiveness of students based on dimensions, the average score between 0 and 5 indicates a passive level, 6 to 11 indicates less assertive, 12 to 18 is moderately assertive. A score between 19 and 24 shows an assertive level, and a score between 25 and 30 indicates an aggressive level (Abdullah, 2005) as shown in Table 4 below.

Table 4. Interpretation of Assertiveness Dimension Score

Score	Level
0-5	Passive
6-11	Less Assertive
12-18	Moderately Assertive
19-24	Assertive
25-30	Aggressive

4.4.2. Social Skills Inventory Del-Prette (SSI Del-Prette)

To measure social skills, the researcher employed the Social Skills Inventory (SSI-Del-Prette), developed by Del Prette and Del Prette (2001; 2011; 2019). In the context of this study, the Malay version of the SSI adapted by Nor Azima (2008) was used. The SSI used in this study consists of 38 items grouped into five factors which is Factor 1: Coping and self-assertion with risk; Factor 2: Self-assertion in the expression of positive affect; Factor 3: Conversation and social confidence; Factor 4: Self-exposure to strangers and new situations; and Factor 5: Self-control of aggressiveness. For each item, respondents were asked to indicate their level of agreement using a four-point Likert scale: 1 = never; 2 = rarely; 3 = almost always; and 4 = always.

The total mean score for overall social skills is 152 for the 38 items. The minimum possible score that can be obtained by a respondent is 38 (1×38), while the maximum possible score is 152 (4×38). For this study, the SSI adapted from Nor Azima's (2008) and the level of students' social skills are classified into three categories, which are low with a mean score from 0 to 50, moderate with a mean score from 51 to 101, and high with a mean score from 102 to 152 as shown in Table 5.

Table 5. Interpretation of Social Skills Score

Score	Level
0-50	Low
51-101	Moderate
102-152	High

Meanwhile, to determine the level of each social skill dimension, a mean score between 0 and 11 indicates that the student's social skill dimension is at a low level, a score between 12 and 24 indicates a moderate level, and a score between 25 and 36 indicates a high level of social skill dimension, as shown in Table 6 below.

Table 6. Interpretation of Social Skills Dimension Score

Score	Level
0-11	Low
12-24	Moderate
25-36	High

4.4.3. Ryff's Psychological Well-Being Scale

The construct of psychological well-being in this study was measured using Ryff's (1989) Psychological Well-Being (PWB) questionnaire, which comprises six dimensions: autonomy, environmental mastery, personal growth, purpose in life, self-acceptance, and positive relations with others. There are three versions of the Ryff instrument: the long version containing 120 items (20 items per dimension) and 84 items (14 items per dimension); the medium version containing 54 items (9 items per dimension); and the short version containing 42 items (7 items per dimension) and 18 items (3 items per dimension) (Ryff, 1989; Ryff & Singer, 1996). The scale has been translated into more than 20 languages worldwide. The 42-item Malay version of the Psychological Well-Being Scale used in this study was adapted from previous studies conducted by Hasimah (2010), Norsayyidatina (2017), Roslan

et al. (2017), and Salina and Rahimi (2015). The six dimensions of well-being were measured using a six-point Likert scale: 1 = strongly disagree; 2 = disagree; 3 = somewhat disagree; 4 = agree; 5 = somewhat agree; and 6 = strongly agree. Negatively worded items were reverse scored. For the 42-item Ryff Psychological Well-Being Scale, the total possible score is 252. The minimum score that can be obtained by a respondent is 42 (1×42), while the maximum score is 252 (6×42).

Based on Norsayyidatina (2017), students' levels of psychological well-being are classified into three categories which are low (0-83), moderate (84-168), and high (169-252) as presented in Table 7 below. High scores across the six dimensions of psychological well-being (PWB) indicate an individual with a high level of well-being, characterized by self-acceptance, personal growth, purpose, effective environmental management, autonomy, and positive relationships (Ryff, 1989; Ryff & Keyes, 1995). Moderate scores suggest that while the individual has areas of psychological strength, they may struggle in one or more dimensions, indicating room for improvement (Ryff, 1989; Ryff & Keyes, 1995). Conversely, low scores across one or more dimensions may reflect poor psychological well-being, suggesting the need for therapeutic intervention or support to enhance their overall well-being (Ryff, 1989; Ryff & Keyes, 1995).

Table 7. Interpretation of Psychological Well-Being Skills Score

Score	Level
0-83	Low
84-168	Moderate
169-252	High

Meanwhile, to determine the level of the students' psychological well-being dimensions, a mean score between 0 and 13 indicates a low level of psychological well-being, 14 to 28 indicates a moderate level, and a mean score between 29 and 42 indicates a high level of psychological well-being, as detailed in Table 8 below

Table 8. Interpretation of Psychological Well-Being Dimension Score

Score	Level
0-13	Low
14-28	Moderate
29-42	High

4.5. Instrument Validity

Validity is an important concept in the context of measuring an instrument. Validity refers to the ability to measure what is intended to be measured (Azizi Yahya et al., Field, 2018). The validity of a measurement also depends on how successfully it measures what it is supposed to measure (Azizi Yahya et al., 2017). Therefore, the validation process should be carried out correctly and precisely. In general, validity can be divided into several types which are face validity, content validity, criterion validity, and construct, each with different purposes and objectives (Taherdoost, 2016). Content validity is the first level of validation that must be conducted before the study and the actual research is carried out on the population. Content validity also refers to evaluating each item in the instrument to ensure it is suitable for the instrument's intended purpose (Mohammad Rahim et al., 2018).

In the context of this study, all three instruments were administered in Malay and adapted from previously validated versions. The Rathus Assertiveness Schedule (RAS) and Ryff's Psychological Well-Being Scale have demonstrated acceptable validity and reliability in prior Malaysian studies. For example, a Malaysian validation study of the RAS conducted by Jusoh et al. (2023) employed exploratory factor analysis and confirmed the factor structure within a local sample, with acceptable reliability (Cronbach's $\alpha = .748$), supporting its use in the Malaysian context. Similarly, the

psychometric properties of Ryff's Psychological Well-Being Scale have been examined in Malaysian samples using factor analytic techniques. For instance, Shariff and Wan Sulaiman (2018) reported that factor analysis of the scale produced a coherent structure with acceptable factor loadings and reliability coefficients across dimensions, indicating that the instrument possesses adequate construct validity within the Malaysian context.

In addition to construct validity evidence from prior studies, content and face validity of the Malay versions were further supported through expert evaluation. Previous studies have reported high Content Validity Index (CVI) values for the Ryff scale (e.g., 0.94 reported by Salina & Rahimi, 2015), indicating strong agreement among experts. In the present study, the instruments were reviewed by experts to ensure linguistic clarity and contextual appropriateness for first-year undergraduate students. However, it is acknowledged that construct validity (e.g., factor structure through CFA or EFA) was not re-examined within the current multi-institution sample. This is consistent with the descriptive nature of the study, which focuses on assessing levels of psychosocial constructs rather than instrument validation. Future research is recommended to conduct confirmatory analyses to further establish the structural validity of these instruments in diverse student populations. Therefore, the researcher performed content validity for the SSI Del-Prette (Del Prette & Del Prette, 2001, 2011, 2019; Nor Azima, 2008).

4.5.1. Process of Determining Content Validity for SSI Del-Prette

For the process of determining the content validity of the SSI, several methods can be used. However, in the context of this study, the researcher chose to use the Content Validity Index (CVI). This choice was based on several advantages of the CVI method compared to other methods. The CVI method was introduced by Waltz and Bausell (1981) and developed by Lynn (1986), and it has been proven to be more practical in terms of time and cost (Tojib & Sugianto, 2006) because it measures the content validity of items through empirical measurement, as compared to other content validity methods that are difficult to interpret. In this study, the CVI method was considered appropriate because it measures the agreement on the relevance of an item within the instrument, as assessed by expert panels. Involvement of a large number of expert panels will provide detailed feedback on the items in the questionnaire. This is in contrast to other methods, such as Cohen's Kappa, Tinsley-Weiss T Index, and James, Demaree & Wolf Index.

Furthermore, the CVI method provides a table for determining the critical value that needs to be achieved (Wilson et al., 2012). The CVI method has also been widely used by local and international researchers to determine the content validity coefficient of instruments (Mohd Effendi et al., 2017). However, the CVI technique is rarely used in social science research, particularly in counselling. The CVI value can be determined using two methods which are the Item Content Validation Index (I-CVI) and the Sum of Content Validation Index (S-CVI). The I-CVI guides the experts to review, reject, or replace the items in the instrument with new ones (Polit & Beck, 2006; Salbiah @ Salleh et al., 2017). In determining the content validity of the SSI-Del-Prette, the researcher used external reviewers, namely the evaluations and critiques from a panel of experts. This expert evaluation method has also been conducted by other researchers, including Aslina (2013), Kamaruzaman Moidunny (2012), Mohd Noor (2012), and Zainol Abidin Ishak (2016).

The selection of experts is crucial in determining the quality of the items and the content of the produced instrument. Salbiah @ Salleh et al. (2017) explain that the criteria for selecting an expert in a field include i) having experience and working for an extended period in the field, ii) holding a PhD degree or qualifications in the field, iii) consistency between experts, meaning an expert with consistent opinions with others, iv) the ability to distinguish deficiencies or changes in an item, and v) attending workshops or training sessions, either in-service or external. Table 3.26 below lists the experts involved in the content validity of the SSI Del-Prette.

According to Najib (1999), for content validity of an instrument, six to nine experts are sufficient for the assessment. Meanwhile, Sousa & Rojjanasrirat (2011) suggest that the expert panel for content validity should consist of six to ten experts. Lynn (1986) also suggests a panel of five to ten experts. Therefore, the researcher appointed six experts, all with academic qualifications at the PhD level and specific expertise relevant to evaluating the SSI questionnaire, including expertise in psychometrics, instrument construction, counseling and psychology modules, and professionals from various higher education institutions (IPTA) and the Public Services Department. The inclusion of experts from various universities and agencies is to avoid bias.

In terms of instrument evaluation by experts, this study used a set of questionnaires built according to Lawshe (1975). The questionnaire used had two parts. The first part was the content validity review, which contained 38 items from five constructs related to social skills, adapted from previous research (Nor Azima, 2008). The five constructs used in this social skills measurement instrument were: coping and self-assertion with risk (10 items), self-assertion in expressing positive effects (8 items), conversation and social confidence (7 items), self-exposure to strangers and new situations (7 items), and self-control of aggressiveness (6 items). Each expert reviewed all items by assessing the relevance of each item based on a 10-point Likert scale, where 1 = very unsuitable and 10 = very suitable.

The second part consisted of a qualitative feedback section, providing the experts with the opportunity to make corrections or suggest more appropriate items in relation to language, society, and culture in Malaysia, or to give written comments and suggestions for suitable statements. Aspects that needed to be considered by the experts during the content validity review procedure of this study included the use of language and statements, language aspects, and ensuring the instrument's suitability for use in higher education contexts in Malaysia. All 38 items were reviewed by the experts to meet the content validity requirements of the SSI. The instrument was emailed to the experts before setting an appointment. For this assessment, the expert panel was given up to two weeks to complete the questionnaire.

The content validity evaluation of the SSI questionnaire in this study was conducted through a systematic process based on recognized procedures in the field of research. The expert panel was asked to evaluate each item using a 10-point scale, from 1 (very unsuitable) to 10 (very suitable). According to Polit & Beck (2006), this scale was converted into a dichotomous form, where scores from 1 to 5 were considered "0" (not agreed/relevant), and scores from 6 to 10 were considered "1" (agreed/relevant). After receiving the expert evaluations, content validity was measured using several key indices, namely the Content Validity Ratio (CVR), which empirically measures how important the items are as considered by the expert panel. The CVI formula is as follows: $CVI = (ne - N/2) / (N/2)$

Where "ne" is the number of experts who rated the item as "important," and N is the total number of experts. The CVI value ranges from -1 to +1, with a value approaching +1 indicating strong agreement that the item is important (Matoré et al., 2017; Lawshe, 1975). Items with a CVI value below 0.50 are usually considered less valid and are either replaced or removed (Mohamed et al., 2017). The Content Validity Index (CVI) is calculated as the average CVI value for all items, which indicates the overall content validity of the instrument. For new instruments, the recommended minimum CVI value is above 0.80 to ensure that the items are clear, relevant, and have high content validity (Lau et al., 2018; Shrotryia & Dhanda, 2019; Polit et al., 2007).

The Sum of Content Validity Index (S-CVI) refers to the overall content validity index based on the expert panel's overall agreement. In this study, the SSI instrument obtained an S-CVI value of 0.99 based on evaluations from six expert panel members, indicating a very high level of content validity and acceptance. This confirms that the items in the instrument are appropriate and accurately represent the construct being measured. In conclusion, the process of content validity evaluation of the SSI instrument using CVR, CVI, and S-CVI is crucial to ensure that the instrument is accurate and appropriate for measuring the constructs being studied.

4.6. Instrument Reliability

According to Hamidah Yusof et al. (2015), to ensure that the research tool used is stable, predictable, and capable of producing accurate and unquestionable findings, reliability analysis is essential because it is related to the concept of validity. Reliability refers to the consistency or stability of the assessment results (Azizi Ahmad, 2010). This means an individual will receive the same score from an instrument, even if measured repeatedly with the same instrument. Generally, reliability can be measured using several methods, such as Kuder Richardson, Internal Consistency, Test-Retest, Inter-Rater Reliability, and Split-Half. However, the method most commonly used by researchers is Cronbach's Alpha reliability coefficient (George & Mallery, 2016).

Therefore, to determine the reliability of all three instruments, the researcher chose to use Cronbach's Alpha method. Cronbach's Alpha is a reliability coefficient used to demonstrate internal consistency or uniformity of the items in the instrument (Sidek Mohd Noh, 1998). It also indicates the strength of elements within a set of questions in relation to one another. A good reliability value for an instrument is an Alpha value greater than 0.6 at the 0.05 significance level (Cresswell, 2016; 2010; Pallant, 2010). However, according to Hair et al. (2014), the Cronbach's Alpha reliability coefficient ranges from 0.00 to 1.00, and the lowest acceptable value is 0.70. Meanwhile, according to Mohd Majid (2005), this value can be as low as 0.60 in exploratory studies.

Although the reliability of the instruments used in this study have been tested and discussed in previous research group discussions, they need to be tested again within the context of Malaysian society. Furthermore, certain modifications in terms of phrasing, place, context, and study subjects can change the dimensions of the instrument's reliability. Therefore, before the actual study was conducted, the instrument was tested through a pilot study held at a university in the eastern coast. In this study, the researcher conducted a pilot study with 137 first year undergraduate's students from the 2021/2022 academic session, using the RAS, SSI Del-Prette, and Ryff's Psychological Well-Being Scales, all in Malay, adapted from previous studies.

Before conducting the pilot study, permission was obtained from the Deputy Vice Chancellor for Student Affairs & Alumni of the relevant university. The questionnaire was administered using Google Forms with the assistance of the university's Student Affairs staff. First-year undergraduate students, randomly selected by the Student Affairs officer, were provided with a link to complete the survey. Students could choose to respond using either a laptop or a smartphone. This method was selected to facilitate data collection and to minimise the likelihood of incomplete responses through the use of required-response settings. A standard set of instructions regarding how to answer the questionnaire, the purpose of the study, and its potential benefits was explained to the participants prior to data collection. Each student was asked to read the instructions carefully and complete the questionnaire independently, based on their own understanding, without assistance from others. The questionnaire consisted of four sections: Section A covered respondents' demographic profiles, Section B assessed assertiveness, Section C measured social skills, and Section D examined psychological well-being.

4.6.1. Reliability of RAS

Several researchers have conducted pilot studies to measure the reliability of the RAS instrument. For example, Abdullah (2005) used the RAS to measure assertive behavior among workers at the Johor State Customs Department. The reliability value obtained in Abdullah's (2005) study was 0.70, indicating that the reliability of the questionnaire used in Abdullah's study was acceptable. In addition, studies by Haladin et al. (2013) and Zainol Abidin (2016) to measure the reliability of this instrument also showed high Cronbach's alpha values, ranging between 0.82 and 0.84. Hawa (2014) also conducted a pilot study to measure the reliability of the RAS and obtained a high Cronbach's alpha value of 0.84.

Since the instrument had already been translated by Abdullah (2010) and Jusah et al. (2023) into the Malay version, the researcher only reviewed the original and translated versions with the help of two language experts who were also skilled in counseling and psychology. However, in this study, due to

the researcher having a different sample of participant which are first year undergraduate student, another pilot study was conducted to test the reliability of the. The overall Cronbach alpha reliability coefficient for the RAS instrument also showed a high value of 0.93. Therefore, this instrument can be used in the actual study to measure assertiveness, as it has been verified to have high consistency and reliability. The detailed reliability values for the RAS are shown in Table 9.

Table 9. Reliability for the RAS Instrument

Construct	Cronbach Alpha (α)
Right to mutual respect	0.87
Right to mutual feelings	0.86
Right to make mistakes	0.85
Right to say "no"	0.88
Right to ask and communicate	0.84
Overall	0.93

4.6.2. Reliability of SSI Del-Prette

In the context of this study, the SSI-Del Prette used was in the Malay version, adapted from Nor Azima's study (2008). Nor Azima (2008) conducted a pilot study on teachers and obtained a high Cronbach's alpha reliability coefficient of 0.79. However, since this study focuses on the social skills of university students, the researcher made modifications, including changes in phrasing, sentence structure, and corrections for technical errors like spelling and language. All 38 items were maintained as in the original instrument. Subsequently, the researcher conducted a pilot study to test the reliability of the SSI Del-Prette on 137 first year undergraduate students for the 2021/2022 academic session at a university on the Eastern Coast.

Based on the pilot study, the Cronbach alpha values range from 0.83 to 0.85. The overall Cronbach alpha reliability coefficient for the instrument is 0.94. The detailed reliability values for the SSI are shown in Table 10. This shows that the SSI instrument has very high reliability. Therefore, when compared to all previous findings, including the pilot study, the results indicate that the SSI instrument has high reliability. As such, the SSI can be used in the actual study to measure university students' social skills as it has been verified to have high consistency and reliability.

Table 10. Reliability for the SSI

Construct	Cronbach Alpha (α)
Coping and self-assertion with risk	0.84
Self-assertion in the expression of positive effect	0.85
Conversation and social confidence	0.84
Self-exposure to strangers and new situations	0.83
Self-control of aggressiveness	0.85
Overall	0.94

4.6.3. Reliability of Ryff's Psychological Well-Being Scale

The Ryff's Psychological Well-Being Scale is one of the most popular research instruments that has been used by many researchers locally to measure psychological well-being across various populations. Several researchers have also conducted pilot studies to obtain the reliability coefficient for this instrument. Among them is Salina & Rahimi Che (2015), who conducted a study involving 210 special education teachers in the Integrated Special Education Program (PPKI) in the states of Penang, Kedah, and Perlis. The results showed that Ryff's Well-Being Scale had a high internal consistency value of Cronbach's alpha of 0.97. Furthermore, Norsayyiditina (2017) also conducted a pilot study and obtained a Cronbach alpha value of 0.70 for the reliability of the Ryff's Psychological Well-Being Scale.

However, since the researcher had a different sample of participants, specifically first-year undergraduate students, the researcher decided to conduct another pilot study to test the reliability of the Ryff's Psychological Well-Being Scale on 137 first year undergraduate students from the 2021/2022 academic session at a public university on the eastern coast. Based on Table 11, the Cronbach alpha reliability value for the entire Ryff's Psychological Well-Being Scale is 0.88. Therefore, overall, based on the results of this pilot study, it is evident that the instruments used in this study have high reliability, and all items in the instruments are suitable for use in the actual study to measure university students' psychological well-being.

Table 11. Reliability for the Ryff's Psychological Well-Being

Construct	Cronbach Alpha (α)
Autonomy	0.85
Environmental mastery	0.83
Personal growth	0.83
Positive relationships	0.84
Purpose in life	0.85
Self-acceptance	0.84
Overall	0.88

5. Findings and Discussion

5.1. Respondent Profile

Although the required sample size was 1,180, but a total of 1,206 first-year undergraduate students from Malaysian public universities participated in this survey study. The profiles of the respondents were described based on key demographic factors, including gender, ethnicity, field of study, place of residence, and socioeconomic status (SES), as detailed in Table 12.

Table 12. Respondent Profile

	Frequency (<i>f</i>)	Percentage (%)
Gender		
Female	702	58.2
Male	504	41.8
Ethnicity		
Malay	678	56.2
Chinese	180	14.9
Indian	110	9.1
Bumiputera Sabah	125	10.4
Bumiputera Sarawak	113	9.4
Field of Study		
Science and Technology	704	58.4
Social Sciences	502	41.6
Place of Residence		
Urban	634	52.6
Rural	572	47.4
Household Income		
Low ($\leq$ RM4,850)	562	47.0
Middle (RM4,851–10,959)	544	45.0
High ($\geq$ RM10,960)	100	8.0
Total	1206	100

The majority of respondents were female students (58.2%), while male students accounted for 41.8%. In terms of ethnicity, most respondents were Malay (56.2%), followed by Chinese (14.9%), Bumiputera Sabah (10.4%), Bumiputera Sarawak (9.4%), and Indian (9.1%). With respect to field of study, a greater proportion of students were enrolled in Science and Technology programs (58.4%), while 41.6% came from the Social Sciences. In terms of residence, slightly more than half of the respondents reported living in urban areas (52.6%), while the remainder were from rural areas (47.4%). Finally, the distribution of household income showed that nearly half of the students were from low-income families (47.0%), followed by middle-income families (45.0%), and a small proportion from high-income families (8.0%). This demographic distribution reflects the diversity of the respondents and enhances the representativeness and credibility of the study's findings.

5.2. Level of Assertiveness among First-Year Undergraduate Students at Public Universities in Malaysia

Table 13 shows the assertiveness levels among 1,206 first year undergraduate students in UA Malaysia. Of this total, 56.60% of students have a moderate level of assertiveness, 16.20% are less assertive, 25.10% have a high level of assertiveness, and 2.10% are categorized as aggressive.

Table 13. Level of Assertiveness Among First-Year Undergraduate Student in UA Malaysia (N=1206)

Score	Level	Frequency (<i>f</i>)	Percentage (%)
0-29	Passive	0	0
30-59	Less Assertive	195	16.20
60-90	Moderately Assertive	682	56.60
91-120	Assertive	303	25.10
121-150	Aggressive	26	2.10

The mean scores for students based on dimensions are shown in Table 14. According to the table, five dimensions of assertiveness were examined. The overall mean assertiveness score for students was ($M = 79.87$, $SD = 18.67$), categorized as moderate. Additionally, all assertiveness dimensions also fall within the moderate level, with mean scores for the right to mutual respect (13.36), right to mutual feelings (18.45), right to make mistakes (15.90), right to say "no" (16.02), and right to ask and communicate (16.15).

Table 14. Mean Score for Assertiveness (N=1206)

Assertiveness Constructs	Mean (<i>M</i>)	Standard Deviation (<i>SD</i>)	Level
Right to mutual respect	13.36	2.98	Moderate
Right to mutual feelings	18.45	4.94	Moderate
Right to make mistakes	15.90	4.16	Moderate
Right to say "no"	16.02	3.95	Moderate
Right to ask and communicate	16.15	4.48	Moderate
Overall	79.87	18.67	Moderate

As shown in Table 14, the overall mean score for assertiveness was $M = 79.87$, placing students at a moderate level of assertiveness. All five dimensions which is respect for others ($M = 13.36$), expressing feelings ($M = 18.45$), acknowledging mistakes ($M = 15.90$), ability to say "no" ($M = 16.02$), and interaction and communication ($M = 16.15$) were also found to be at moderate levels. Assertiveness is a critical skill for effective communication, personal development, and mental health, particularly for university students (ElBarazi et al., 2024; Samfira, 2022). However, the findings of this study revealed that the overall level of assertiveness among first-year undergraduates in Malaysian public universities was at a moderate level. This finding is consistent with previous studies conducted in Malaysia. For example, Salleh and Zuria (2007) reported that assertiveness, measured using the Rathus Assertiveness

Schedule (RAS), was closely associated with life satisfaction and effective communication skills among university students.

The RAS itself measures the extent to which individuals can express opinions and emotions confidently without offending others (Rathus, 1973). Thus, the moderate levels observed in this study suggest that students are still in the process of developing essential assertiveness skills that are crucial for interpersonal communication and psychological well-being. Similarly, Haladin et al. (2013) found that assertiveness was strongly linked with life satisfaction, active learning engagement, and social interaction. Cultural factors may also explain the moderate levels of assertiveness among Malaysian undergraduates. As Malaysia is a collectivist society that emphasizes social harmony and respect for authority, students may be less likely to assert themselves compared to their Western peers, who are often encouraged to adopt individual assertiveness (Abdul Majid et al., 2024; Haladin et al., 2013). Ang (2016) and Sieng and Yusof (2017) likewise highlighted that Malaysian students often lack the confidence to express and defend their opinions constructively.

A study conducted at Manipal University College Malaysia (MUCM) further revealed that the majority of medical students displayed tentative behaviors (77.5%), while only 21.7% were identified as assertive (Alyssa Sureyya et al., 2021). This underscores the tendency for Malaysian students to be less assertive than their international peers. Barriers such as fear of making mistakes, cultural emphasis on deference, and limited opportunities to practice assertive communication further contribute to this trend. Students may avoid speaking up due to anxiety about being judged or ridiculed, particularly when using English as a second language. This avoidance is reinforced by an education system that prioritizes technical knowledge over interpersonal and communication skills (Haladin et al., 2013; Mohamad et al., 2024). Without practical exposure to assertive communication, graduates may struggle in situations that require confidence and self-expression.

Low levels of assertiveness can negatively impact academic performance, interpersonal relationships, and employability. In academic settings, non-assertive students may hesitate to express their ideas, negotiate roles in group assignments, or seek clarification, leading to weaker collaboration and lower performance (Blegur et al., 2023; Toktaş et al., 2022). Studies have also shown a negative correlation between assertiveness and attitudes toward academic courses, with less assertive students often displaying lower motivation and reduced achievement (Alyssa Sureyya et al., 2021; Toktaş et al., 2022). Similarly, low assertiveness hinders the ability to build and sustain healthy relationships, as students may struggle to communicate their needs or establish boundaries (Moss et al., 2021).

Research has also demonstrated a link between low assertiveness and higher levels of academic anxiety, which can disrupt concentration, attention, and academic functioning (ElBarazi et al., 2024; Mohebi et al., 2012). In terms of employability, employers often seek candidates who can communicate effectively and advocate for themselves (Haladin et al., 2013; Vargas-Saritama et al., 2025). Graduates with low assertiveness may underperform in job interviews and workplace interactions, which can limit their career prospects. They may also be perceived as lacking confidence or initiative (Alyssa Sureyya et al., 2021), potentially reducing their chances of being hired or promoted. Taken together, these findings highlight the importance of developing assertiveness as a vital skill for university students.

Moderate levels of assertiveness indicate room for targeted interventions and training programs designed to strengthen students' self-confidence and interpersonal abilities, thereby supporting their academic success, psychological well-being, and employability. Importantly, interventions must be culturally sensitive, taking into account the collectivist values of Malaysian society to ensure effectiveness. Universities should prioritize the design and implementation of assertiveness development programs and integrate counselling support services to identify and assist students at risk of low assertiveness early in their studies. By fostering a supportive environment, universities can encourage students to express their thoughts and feelings more confidently and prepare them for future academic and professional challenges.

5.3. Level of Social Skills among First-Year Undergraduate Students at Public Universities in Malaysia

Table 15 below shows the classification of social skills levels among 1,206 first year undergraduate students in UA Malaysia. The majority of students, 64.10%, have a moderate level of social skills, and 12.90% have low social skills. Meanwhile, only 23% of respondents are in the high social skills category.

Table 15. Level of Social Skills Among First-Year Undergraduate Student in UA Malaysia (N=1206)

Score Range	Level	Frequency (<i>f</i>)	Percentage (%)
0-50	Low	156	12.90
51-101	Moderate	773	64.10
102-152	High	272	23

The mean scores and social skills levels based on dimensions are discussed in Table 16. There are five dimensions of social skills being examined. The overall mean score for social skills is ($M = 86.49$, $SD = 24.49$), classified as moderate. Additionally, all dimensions of social skills are also at the moderate level, with mean scores for the dimensions coping and self-assertion with risk (22.60), self-assertion in the expressing of positive effects (19.01), conversation and social confidence (15.64), self-exposure to stranger and new situations (14.92), and self-control from aggressiveness (14.17).

Table 16. Mean Score for Social Skills (N=1206)

Social Skills Constructs	Mean (<i>M</i>)	Standard Deviation (<i>SD</i>)	Level
Coping and self-assertion with risk	22.60	6.87	Moderate
Self-assertion in the expression of positive effect	19.01	6.30	Moderate
Conversation and social confidence	15.64	4.42	Moderate
Self-exposure to strangers and new situations	14.92	5.22	Moderate
Self-control of aggressiveness	14.17	3.08	Moderate
Overall	86.49	24.49	Moderate

The findings of this study indicate that overall, first-year undergraduate students in Malaysian public universities demonstrated a moderate level of social skills. Analysis of the five dimensions which is Coping and self-assertion with risk, self-assertion in the expression of positive effect, conversation and social confidence, self-exposure to strangers and new situations, and self-control of aggressiveness also revealed moderate mean scores across the board. These results align with previous studies that have identified limited social competence among Malaysian undergraduates. Chin (2016) reported that students often display lower levels of communicative competence, assertiveness, and prosocial tendencies, while Nooriah and Zakiah (2017) also found that many lacked strong interpersonal abilities. Similarly, Kipli et al. (2022) observed that students in Sarawak demonstrated social skills at unsatisfactory moderate levels, suggesting that a considerable number of university students in Malaysia still lack sufficient mastery of social competence to navigate daily life effectively.

Low levels of social skills are associated with numerous challenges, including difficulty in forming and maintaining friendships, which may contribute to loneliness and social isolation in the demanding university environment (Moeller & Seehuus, 2019). Poor communication skills can also negatively affect classroom participation, group projects, and the ability to seek academic support (Ayllón-Salas, 2024; Gul et al., 2023; Shengyao et al., 2024; Yan Carlo, 2022). Furthermore, deficits in social competence have been linked to heightened risks of anxiety and depression among students (Rikizia et

al., 2024; Segrin, 2019). Beyond academics, inadequate social skills may also hinder employability. Employers frequently emphasize the importance of interpersonal and communication skills for career success (Suleman et al., 2022; Vanhove et al., 2024).

Graduates with weaker social skills may underperform in interviews, workplace collaboration, and leadership roles, which can negatively affect career progression and job satisfaction. In summary, social skills serve as a foundation for effective communication, interpersonal relationships, academic performance, mental health, and employability. Current higher education policies in Malaysia, such as the Malaysia Education Blueprint 2015–2025, emphasize the development of soft skill including social competence as key to producing competitive and holistic graduates (Ministry of Education Malaysia, 2015). Therefore, universities should prioritize the integration of social skills development into curricula and provide structured training and intervention programs to better prepare students for academic and professional success.

5.4. Level of Psychological Well-Being among First-Year Undergraduate Students at Public Universities in Malaysia

Table 17 below shows that 67.70% of the 1,206 students exhibit a moderate level of psychological well-being, while 14.20% have a low level, and only 18.10% of students report a high level of psychological well-being.

Table 17. Level of Psychological Well-Being Among First-Year Undergraduate Student in UA Malaysia (N=1206)

Score Range	Level	Frequency (<i>f</i>)	Percentage (%)
0-83	Low	171	14.20
84-168	Moderate	817	67.70
169-252	High	218	18.10

The mean scores and levels of psychological well-being among students based on dimensions are discussed in Table 18. The overall mean psychological well-being score for students is ($M = 136.13$, $SD = 38.38$), indicating a moderate level. Additionally, six dimensions of psychological well-being were studied, and all dimensions also fall within the moderate level, with mean scores for the dimensions as follows: autonomy (22.73), environmental mastery (23.22), personal growth (22.30), positive relations with others (22.36), purpose in life (23.05), and self-acceptance (22.46).

Table 18. Mean Score for Psychological Well-Being (N=1206)

Psychological Well-Being Dimension	Mean (<i>M</i>)	Standard Deviation (<i>SD</i>)	Level
Autonomy	22.73	7.24	Moderate
Environmental mastery	23.22	6.28	Moderate
Personal growth	22.30	6.42	Moderate
Positive relationships	22.36	6.60	Moderate
Purpose in life	23.05	6.62	Moderate
Self-acceptance	22.46	6.80	Moderate
Overall	136.13	38.38	Moderate

The study also found that the overall psychological well-being of first-year undergraduates was at a moderate level, with all six-dimension self-acceptance, positive relationships, autonomy, environmental mastery, purpose in life, and personal growth also falling within the moderate range. This finding is consistent with studies in Malaysia, such as Aziz et al. (2024), who reported moderate levels of psychological well-being among students at Universiti Malaysia Kelantan. National survey data also

highlight significant levels of distress among Malaysian youth, with Soo et al. (2024) documenting that 34% of students experienced anxiety, 27.5% faced depression, and 18.6% reported stress. Internationally, similar concerns have been raised, with more than half of university students in the United Kingdom and the United States reporting clinically significant levels of anxiety and depression (Myrissa et al., 2024). The COVID-19 pandemic further exacerbated these issues, intensifying psychological pressures faced by students globally (Wong et al., 2023).

Academic pressures, frequent examinations, and concerns about future employment were identified as major stressors, leading to lower psychological well-being (Chaudhry et al., 2024). Students with lower well-being often experience impaired concentration, reduced problem-solving skills, and weaker coping strategies, placing them at higher risk of depression and anxiety (Wąsowicz et al., 2021). Limited psychological flexibility also restricts their ability to adapt to change and manage negative emotions effectively. Low psychological well-being has further implications for social participation. Students may withdraw from social and co-curricular activities, misinterpret social cues, and feel isolated, which can reduce life satisfaction and hinder personal development (Campbell et al., 2022; Chaudhry et al., 2024). Poor coping mechanisms, such as avoidance or unhealthy behaviors, may worsen stress and negatively affect both mental and physical health (Indreswari et al., 2024). Such challenges increase the risk of academic disengagement and dropout, with long-term consequences for careers and life outcomes. Employability is also influenced by students' psychological well-being.

Research shows that students with poor mental health may struggle to build professional networks and engage effectively in teamwork, both of which are critical for career advancement (He et al., 2022). Feelings of isolation may further reduce collaboration and limit exposure to diverse perspectives in professional contexts. These findings underscore the importance of prioritizing psychological well-being in higher education. National policies, such as the Malaysia Education Blueprint 2015–2025 (Ministry of Education Malaysia, 2015) and the National Mental Health Policy 2020–2025, highlight the need for early intervention and support services for young people, including undergraduates. Universities should therefore strengthen psychosocial support systems, reduce stigma around mental health, and foster inclusive learning environments that encourage students to seek help. Such efforts will not only enhance students' academic and personal development but also prepare them for long-term resilience and career readiness.

6. Conclusion

This study examined the levels of assertiveness, social skills, and psychological well-being among first-year undergraduate students in Malaysian public universities. The findings indicate that students generally demonstrate moderate levels across all three constructs and their respective dimensions. While these results suggest that students possess foundational competencies in self-expression, interpersonal communication, and emotional well-being, these skills may not yet be sufficiently developed to meet the increasingly complex academic, social, and professional demands of higher education.

The moderate level of assertiveness suggests that many students are still developing confidence in expressing their opinions and articulating their ideas, which may be influenced by cultural norms that emphasise social harmony and respect for authority. Similarly, the findings on social skills indicate that students may encounter challenges in establishing effective interpersonal relationships, adapting to new environments, and communicating effectively, all of which are essential for success in academic and professional settings. The moderate level of psychological well-being reflects the ongoing challenges faced by students in adjusting to university life, particularly in balancing academic expectations with personal and social demands.

These findings highlight the importance of strengthening students' psychosocial competencies as part of holistic higher education development. Without adequate institutional and social support, students may face difficulties in sustaining academic engagement, maintaining psychological well-being, and

preparing for future employability. In light of these findings, several recommendations are proposed. First, higher education institutions should integrate soft skills such as assertiveness, communication, and social competence into the formal curriculum to ensure systematic development of these competencies. Second, counselling and psychoeducational programmes should be expanded to include assertiveness training, resilience-building, and emotional regulation strategies, with early identification of students who may require additional support. Third, interventions should be culturally responsive, balancing collectivist values with the need to encourage appropriate self-expression, with an emphasis on group-based and peer-supported approaches.

In addition, universities should strengthen collaboration with industry partners to provide experiential learning opportunities, such as internships, service-learning, and work-based learning, enabling students to apply interpersonal and communication skills in real-world contexts. Efforts should also be made to enhance awareness of mental health and reduce stigma, by fostering supportive and inclusive environments where students feel comfortable expressing their concerns. Finally, institutional strategies should be aligned with national frameworks such as the Malaysia Education Blueprint 2015–2025 and the National Mental Health Policy 2020–2025, which emphasise holistic student development. Overall, enhancing assertiveness, social skills, and psychological well-being among undergraduates requires a structured, culturally responsive, and comprehensive approach that integrates curriculum development, counselling support, experiential learning, and institutional policy alignment.

7. Limitation of Study

This study has several limitations that should be acknowledged. First, the cross-sectional design captures data at a single point in time and therefore does not allow for the examination of changes over time or the establishment of causal relationships among variables. Second, the study relies on self-reported questionnaire data, which may be subject to common-method bias, including social desirability and response consistency effects. Although measures were taken to ensure anonymity and independent responses, such biases cannot be entirely eliminated. Third, the interpretation of results based on categorical thresholds (e.g., low, moderate, and high levels) is dependent on cut-off values adapted from prior studies. Different classification criteria may yield slightly different interpretations of the data. Therefore, the findings should be interpreted with consideration of this threshold sensitivity. Finally, as the sample was drawn from selected public universities, the findings are limited to the sampled institutions and should be interpreted with caution when generalised to all Malaysian public universities.

References

- Abdullah, K. (2005). *Tahap ketegasan pelajar universiti: Satu kajian di Universiti Teknologi Malaysia*. (Doctoral dissertation, Universiti Teknologi Malaysia).
- Abdul Majid, N., Abdullah, N. N., Ramdzan Ali, A. A. E., Ramli, M. S., & Ayob, N. H. (2024). Exploring the interplay between undergraduates' perception of Islamic communication ethics and their practice in assertive communication. *International Journal of Academic Research in Business and Social Sciences*, 14(6), 22–34. <https://doi.org/10.6007/IJARBSS/v14-i6/21542>
- Alberti, R. E., & Emmons, M. L. (2017). *Your perfect right: Assertiveness and equality in your life and relationships* (11th ed.). New Harbinger Publications.
- Alyssa Sureyya, A., Nor, M. M., & Hassan, N. (2021). A survey on assertiveness among undergraduate medical students in Manipal University College Malaysia (MUCM). *Advances in Applied Psychology* 4(2), 237-56.
- Ang, C. S. (2016). Assessment of social competence: exploring the dimensions of young adult social competence. *Jurnal Psikologi Malaysia*, 30(1).
- Bisschoff, Z. S., & Massyn, L. (2025). A conceptual soft skills competency framework for enhancing graduate intern employability. *Higher Education, Skills and Work-Based Learning*, 15(7), 66-81. <https://doi.org/10.1108/HESWBL-08-2023-0239>
- Ayllón-Salas, P., & Fernández-Martín, F. D. (2024). The role of social and emotional skills on adolescents' life satisfaction and academic performance. *Psychology, Society & Education*, 16(1), 49-56. <https://doi.org/10.21071/pse.v16i1.16625>
- Aziz, A. R. A., Hasbullah, N. I., & Shafie, A. A. (2024). Academic stress and psychological wellbeing among university students. *International Journal of Education, Psychology and Counseling*, 9(54), 212-222. <https://doi.org/10.35631/IJEPC.954016>
- Blegur, J., Haq, A. H. B., & Barida, M. (2023). Assertiveness as a new strategy for physical education students to maintain academic performance. *The Qualitative Report*, 28(3), 865-885. <https://doi.org/10.46743/2160-3715/2023.5659>
- Chaudhry, S., Tandon, A., Shinde, S., & Bhattacharya, A. (2024). Student psychological well-being in higher education: The role of internal team environment, institutional, friends and family support and academic engagement. *Plos one*, 19(1), e0297508. <https://doi.org/10.1371/journal.pone.0297508>
- Cho M., Kim O., Pang Y., Kim B., Jeong H., Lee JJ., Jung HH., Jeong S. Y., Park H. Y., Choi H. & Dan H. (2021). Factors affecting frontline Korean nurses' mental health during the COVID-19 pandemic. *International Nursing Review* 68(2), 256-265. <https://doi.org/10.1111/inr.12679>
- Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, 1(3), 98-101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Cohen, L., Manion, L. & Morrison, K. 2018. *Research Methods in Education. Professional Development in Education*. Edisi ke-6. London: Routledge Taylor & Francis Group. <https://doi.org/10.4324/9780203224342>
- Creswell, J. W. 2015. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New Jersey: Pearson.

- ElBarazi, I., Samfira, E. M., & Mohebi, S. (2024). Assertiveness, stress, and well-being in higher education students. *BMC Psychology*, 12(6), 321–335. <https://doi.org/10.1186/s40359-024-01125-7>
- Freitas, P. H. B. D., Meireles, A. L., Ribeiro, I. K. D. S., Abreu, M. N. S., Paula, W. D., & Cardoso, C. S. (2023). Symptoms of depression, anxiety and stress in health students and impact on quality of life. *Revista Latino-Americana De Enfermagem*, 31, e3884. <https://doi.org/10.1590/1518-8345.6315.3885>
- Falsafah Pendidikan Kebangsaan. (2013). Ministry of Education Malaysia.
- Gul, R., Batool, S., Khan, S. I., & Jabeen, F. (2023). The effects of social skills on academic competencies among undergraduate students. *Russian Law Journal*, 11(3S), 308-315.
- Haladin, N. A. B., Ibrahim, N. A., & Rajab, A. (2013). Assertiveness among undergraduates in a Malaysian public university: Implications towards improved employability. *Sains Humanika*, 65(2). <https://doi.org/10.11113/jt.v65.2346>
- Hawa Omar. (2015). *Hubungan antara tingkah laku asertif dengan komitmen organisasi dalam kalangan staf sokongan Universiti Tun Hussein Onn Malaysia (UTHM)* (Master's thesis, Universiti Teknologi Malaysia).
- Indreswari, H., Probawati, D., & Rachmawati, I. (2024). Psychological well-being and student academic burnout. *Jurnal Kajian Bimbingan Dan Konseling*, 7(3), 13. <https://doi.org/10.17977/um001v7i32022p138-149>
- Institute for Public Health. (2015). *National health and morbidity survey 2015: Mental health problems among Malaysian adults*. Ministry of Health Malaysia.
- Institute for Public Health. (2023). *National health and morbidity survey 2023: Key findings*. Ministry of Health Malaysia.
- Jusoh, A. J., Imami, M. K. W., Handrianto, C., Isa, A. N. M., Omar, S. Z., Abdullah, A., & Wahab, S. (2023). Verification of Reliability and Validity of a Malaysian Version of Rathus Assertiveness Schedule as Drug Prevention Scale. *Islamic Guidance and Counseling Journal*, 6(2). <https://doi.org/10.25217/0020236369700>
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610.
- Kipli, M., Rahman, S., & Surat, S. (2022). Kebimbangan dan kemahiran sosial dalam kalangan pelajar di sebuah Institusi Pendidikan Islam di Sarawak. *Jurnal Dunia Pendidikan*, 4(2), 49-62.
- Korda, M., Shulhai, A., Shevchuk, O., Shulhai, O., & Shulhai, A. M. (2025). Psychological well-being and academic performance of Ukrainian medical students under the burden of war: a cross-sectional study. *Frontiers in Public Health*, 12, 1457026.
- Labrague, L. J., & De los Santos, J. A. A. (2020). COVID-19 anxiety among front-line nurses: Predictive role of organisational support, personal resilience and social support. *Journal of Nursing Management*, 28(7), 1653-1661. <https://doi.org/10.1111/jonm.13121>
- Lazarus, A. A. (1971). *Behavior Therapy and Beyond*. New York: McGraw-Hil.
- Lazarus, A. A. (1973). Assertive behavior: A review and theoretical formulation. *Psychological Bulletin* 79(5): 376-391.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.

Mahdavi, P., Valibeygi, A., Moradi, M., & Sadeghi, S. (2023). Relationship between achievement motivation, mental health and academic success in university students. *Community Health Equity Research & Policy*, 43(3), 311-317. <https://doi.org/10.1177/0272684X2110259>.

Ministry of Education. (2024). Higher Education Annual Report 2024. Ministry of Education Malaysia.

Ministry of Education Malaysia. (2008). Falsafah Pendidikan Kebangsaan [National Philosophy of Education]. Ministry of Education Malaysia.

Ministry of Education Malaysia. (2015). Malaysia Education Blueprint 2015–2025 (Higher Education). Ministry of Education Malaysia.

Moeller, R. W., & Seehuus, M. (2019). Loneliness as a mediator for college students' social skills and experiences of depression and anxiety. *Journal of Adolescence*, 73, 1-13. <https://doi.org/10.1016/j.adolescence.2019.03.006>

Mohamad, N. M., Jaaffar, A. H., Abd Majid, N., & Isa, A. (2025). The effect of essential skills on perceived employability: a case of Malaysian undergraduates in energy-focused universities. *Higher Education, Skills and Work-Based Learning*, 15(2), 349-368. <https://doi.org/10.1108/HESWBL-12-2023-0338>

Mohzana, M. (2024). The impact of the new student orientation program on the adaptation process and academic performance. *International Journal of Educational Narratives*, 2(2), 169-178. <https://doi.org/10.70177/ijen.v2i2.763>

Moss, J. G., Firebaugh, C. M., & Morgan, S. M. (2021). Assertiveness, self-esteem, and relationship satisfaction. *International Journal of Arts and Social Science*, 4(2), 235-245.

Myrissa, K., Court, C., & Kelaiditi, E. (2024). Cross-sectional study examining the association between diet quality and the prevalence of anxiety and depression in UK undergraduate students. *Nutrition Bulletin*, 49(3), 383-395. <https://doi.org/10.1111/nbu.12694>

Neuman, W.L. (2009). Social research methods: *Qualitative and quantitative approaches* (7th ed.). Boston, MA: Pearson/Allyn & Bacon.

Norsayyidatina Che Rozubi. (2017). *Kesan program kesejahteraan psikologi menggunakan kaedah bermain kepada kanak-kanak sekolah* (Doctoral dissertation, Universiti Malaya).

Parray, W. M., Kumar, S., David, B. E., & Khare, S. (2020). Assertiveness predicts self-esteem, academic achievement, and stress: A study of Kashmiri adolescents. *Humanities & Social Sciences Reviews*, 8(1), 707-715. <https://doi.org/10.18510/hssr.2020.8185>

Rathus, S. A. (1973). A 30-item schedule for assessing assertive behavior. *Behavior Therapy*, 4(3), 398–406. [https://doi.org/10.1016/S0005-7894\(73\)80120-0](https://doi.org/10.1016/S0005-7894(73)80120-0)

Ranita, H., Sohana, I., & Marhaini, M. (2019). Depression, anxiety, and stress among Malaysian undergraduates: A survey study. *Malaysian Journal of Psychiatry*, 28(1), 23–32.

Rizkia Mulyana, F., Suherman, A., Juliantine, T., Blegur, J., Rohyana, A., & Saeful Bakhri, R. (2024). Building mental health and social skills: The positive impact of jigsaw model in taekwondo course. *Retos: Nuevas Perspectivas de Educación Física, Deporte y Recreación*, (61). <https://doi.org/10.47197/retos.v61.108372>

Rosenfeld, R. B. A. L. B., & Proctor, R. F. (2000). *Interplay: The Process of Interpersonal Communication*. New York: Harcourt Brace.

Ruggeri, K., Garcia-Garzon, E., Maguire, Á., Matz, S., & Huppert, F. A. (2020). Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries. *Health and quality of life outcomes*, 18(1), 192. <https://doi.org/10.1186/s12955-020-01423-y>.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>

Ryff, C. D. (2018). Well-being with soul: Science in pursuit of human potential. *Perspectives on Psychological Science*, 13(2), 242–248. <https://doi.org/10.1177/1745691617699836>

Salina, M. N., & Rahimi, N. M. (2015). Hubungan antara sokongan sosial dan kesejahteraan psikologi dalam kalangan pelajar universiti. *Jurnal Psikologi Malaysia*, 29(1), 1–12.

Salleh, A., & Zuria, M. (2007). Assertiveness and communication skills among Malaysian higher education students. *Journal of Educational Psychology and Counselling*, 1(2), 35–49.

Samfira, E. M. (2022). Differences in the assertiveness level of First Year Pre-Service Teachers from a Romanian University of Life Sciences. *Journal Plus Education/Educația Plus*, 30(1).

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons.

Sekaran, U. (2005). *Research Methods for Business: A Skill Building Approach*. New Jersey: John Wiley and Sons

Segrin, C. (2019). Indirect effects of social skills on health through stress and loneliness. *Health communication*, 34(1), 118-124.

Shariff, M. I. A. S., & Sulaiman, W. S. W. (2018). Penerokaan ciri-ciri psikometrik skala kesejahteraan psikologi dalam kalangan sampel Malaysia (Exploring the psychometric properties of the Psychological Well-being Scale among Malaysian samples). *Jurnal Psikologi Malaysia*, 32(3).

Shengyao, Y., Xuefen, L., Jenatabadi, H. S., Samsudin, N., Chunchun, K., & Ishak, Z. (2024). Emotional intelligence impact on academic achievement and psychological well-being among university students: the mediating role of positive psychological characteristics. *BMC psychology*, 12(1), 389. <https://doi.org/10.1186/s40359-024-01886-4>

Sial, Z. A., Naz, F. L., & Rasheed, A. (2021). Relationship Between Students' Social Skills and Academic Achievement at University Level. *VFAST Transactions on Education and Social Sciences*, 9(3), 232-241.

Sieng, L. W., & Yussof, I. (2017). Comparative study of Malaysia human capital with selected ASEAN and Developed Countries: A fuzzy TOPSIS method. *Geografia-Malaysian Journal of Society and Space*, 11(6), 11-22.

Slimmen, S., Timmermans, O., Mikolajczak-Degrauwe, K., & Oenema, A. (2022). How stress-related factors affect mental wellbeing of university students A cross-sectional study to explore the associations between stressors, perceived stress, and mental wellbeing. *PloS one*, 17(11), e0275925. <https://doi.org/10.1371/journal.pone.0275925>

- Soo, Y. Y., Wong, Y. Y., Ong, S. C., & Ooi, G. S. (2024). Perceptions and beliefs towards mental health and mental illness: A qualitative study among university students in Malaysia. *Malaysian Journal of Medicine and Health Sciences*, 20(1), 70-77. <https://doi.org/10.47836/mjmhs.20.1.10>
- Suleman, N., Admani, A., Rahima, R., Ali, S. S., & Sami, A. (2022). *How do skills influence the students' employability in a developing economy?*
- Suvera, P. (2016). Psychological well-being: A comparative study of tribal and non-tribal college students. *Indian Journal of Health & Wellbeing*, 7(3).
- Tang, Y. Y., Tang, R., & Gross, J. J. (2019). Promoting psychological well-being through an evidence-based mindfulness training program. *Frontiers in human neuroscience*, 13, 237. <https://doi.org/10.3389/fnhum.2019.00237>.
- Tobin, S. S., & Neugarten, B. L. (1961). Life satisfaction and social interaction in the aging. *Journal of Gerontology*. <https://doi.org/10.1093/geronj/16.4.344>
- Toktas, S., Demir, A., & Barut, A. İ. (2022). Examining of University Students' Assertiveness Levels and Attitudes towards Physical Education Course. *African Educational Research Journal*, 10(3), 388-393. <https://doi.org/10.30918/AERJ.104.21.150>.
- Vanhove, A., Opdecam, E., & Haerens, L. (2024). Fostering social skills in the Flemish secondary accounting education: perceived challenges, opportunities, and future directions. *Accounting Education*, 33(4), 414-449. <https://doi.org/10.1080/09639284.2023.2208106>.
- Vargas-Saritama, A., Espinoza-Celi, V. S., & Carrión-Robles, F. (2025). The interrelationship between 21st century skills that enhance graduate employability. *Higher Education, Skills and Work-Based Learning*, 15(2), 403-420. <https://doi.org/10.1108/HESWBL-01-2024-0012>.
- Wąsowicz, G., Mizak, S., Krawiec, J., & Białaszek, W. (2021). Mental health, well-being, and psychological flexibility in the stressful times of the COVID-19 pandemic. *Frontiers in psychology*, 12, 647975. <https://doi.org/10.3389/fpsyg.2021.647975>
- Wong, S. S., Wong, C. C., Ng, K. W., Bostanudin, M. F., & Tan, S. F. (2023). Depression, anxiety, and stress among university students in Selangor, Malaysia during COVID-19 pandemics. *PLOS ONE*, 18(1), e0280680. <https://doi.org/10.1371/journal.pone.0280680>
- World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum.
- Xiang, Y., Shuai, C., Zhang, Y., & Li, Y. (2024). The mental health of college students and the level of academic achievement: Knowledge about psychological well-being, self-control, and college learning. *Československá psychologie*, 68(1), 80-95. <https://doi.org/10.51561/cspsych.68.1.80>
- Yan Carlo, Q. Q. (2022). Social Skills and academic performance in Educational Institutions of High Andean Puno-Peru in 2020. *Journal of Positive School Psychology*, 6(8).
- Zainol Abidin Ishak. (2016). *Kesan modul bimbingan kaunseling kelompok penyesuaian remaja terhadap motivasi, ketegasan, empati, harga diri dan kecerdasan emosi dalam kalangan pelajar salah laku disiplin* (Doctoral dissertation, Universiti Pendidikan Sultan Idris).

Exploring The Professional Development Needs of Early-Career EFL Teachers in Uzbekistan: Implications for University-Based Teacher Training Programs

ABSTRACT

Authors: ¹Munisa G. Muhtarova and
^{1*}Tajularipin Sulaiman

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:

tajulas@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

The research aims to investigate the professional development (PD) needs of early-career English as a Foreign Language (EFL) teachers in public schools of Olmazor District, Tashkent, Uzbekistan. An exploratory-descriptive qualitative research approach was adopted for the study. A context-demographics questionnaire and five focus group discussions with 25 randomly selected teachers meeting the criteria of CEFR B2+ and less than five years of teaching experience were conducted. Five areas of professional development need of early-career EFL teachers were identified: managing large mixed-ability classes, teaching under technology constraints, teachers' speaking confidence and communicative classroom practice, the practical and context-relevance of PD, and follow-up support for teachers after PD. These findings were then translated into a five-module needs-based PD prototype with a focus on differentiation, minimum viable technology, microteaching, communicative lesson redesign, and transfer support through peer observation and professional learning community engagement. This study can inform ministry stakeholders, university-based PD units, and development partners seeking to improve EFL teacher capacity in Uzbekistan

Keywords: Teacher Professional Development, Needs Analysis, EFL Teachers, Early-Career Teachers, Focus Group Discussion, Uzbekistan.

1. Introduction

In recent years, Uzbekistan has made significant strides in education reform, particularly in foreign language education (Abdullaev, 2021; Eshchanova et al., 2020; Presidential Decree PQ-5117, 2021; Presidential Decree PF-79, 2023). As the global standing of English has gone up, the Uzbekistan government has given special focus to developing the capability of English as a Foreign Language (EFL) teachers of the country's public schools (Presidential Decree PQ-4963, 2021). In this context, there has been a significant shift in educational agendas, with an emphasis on equipping learners with the language competencies necessary to thrive in the modern, globalized world (Abdullaev, 2021; Eshchanova et al., 2020). However, despite these developments, in-service EFL teachers continue to face a myriad of challenges that compromise their ability to deliver quality language teaching (Ibragimov & Muradov, 2022; Alkhateeb et al., 2022).

Most pressing, probably, is the very large size of most classes, well beyond the optimum for interactive and personalized instruction. In addition, the lack of modern instructional equipment and technology is part of the circumstance in which instructors cannot adequately motivate students and are limited in how they might use innovative approaches to instruction. On top of these are the often inadequate supports teachers have in the way of professional development (Mamadaminova & Madalińska-Michalak, 2025). While there is growing awareness of the requirement of ongoing professional education for teachers, many teachers still report that professional development activities provided to them are not addressed to their individual situations or are incompatible with the needs of their classrooms (Mamadaminova & Khadjikhanova, 2023).

The in-service EFL teachers' professional development requirements in Uzbekistan are important in the sense that there is an urgent call for high school teacher development programs with precise addresses to the overall educational objectives and specific concerns of teachers in their daily practice (Mamadaminova & Madalińska-Michalak, 2025). This study is specifically concerned with the professional development needs of the in-service EFL teachers in Uzbekistan by using a holistic needs analysis of the teachers of the Olmazor District of Tashkent, Uzbekistan. The study adopts a qualitative research design with the aim of discovering areas of priority concern and determining the nature of support the teachers require to improve their pedagogical competence and teaching practice. The findings of this study will provide valuable information on the most important teacher professional development needs, for example, additional classroom management, additional confidence in speaking, teaching with technology, and additional communication with students.

Furthermore, this study seeks to highlight the necessity of going beyond traditional, generic training programs and into more contextualized efforts that directly address teachers' real needs in Uzbekistan classrooms. By offering recommendations for creating teacher development programs that are efficient, significant, and pertinent to the needs of teachers, this study aims to play a part in the improved study of EFL in Uzbekistan and the professional growth of in-service teachers. This study is intended to inform ministry officials, university-based professional development (PD) units, and donor/partner organizations involved in strengthening EFL teacher capacity in Uzbekistan.

Although enhancing instruction depends on teachers developing professionally, the in-service EFL teachers at many public schools make complaints about the inadequacy of practicality,

misalignment with everyday classroom realities, and a lack of follow-up in many PD options. In Uzbekistan, such concerns have had more influence on early-career teachers struggling to handle mixed proficiency levels, large classes, and a general imbalance in the distribution of materials, technologies, and other forms of teaching support. The situation calls for a grounded and context-sensitive needs analysis that determines what teachers report the most challenges with, what types of PD formats feel genuinely useful to them, and what design decisions can help the teacher apply knowledge in the classroom. The current study, therefore, adopts an exploratory-descriptive approach to investigate the professional development needs of early-career in-service EFL teachers. It seeks to establish a well-evidenced basis for a workable professional development framework suited to the local context. Its findings are expected to help the officials of the Ministry, University Professional Development Units, and donor partners working towards teacher professional development and enhancement of classroom practice in Uzbekistan.

Against this background, the present study investigates the professional development needs of early-career in-service EFL teachers in public schools in Olmazor District, Tashkent, Uzbekistan. Using a context-demographics questionnaire and semi-structured focus group discussions, the study identifies teachers' priority professional development needs, examines perceived gaps in existing PD provision, and translates these findings into a context-sensitive PD prototype. By focusing on early-career teachers working in resource-constrained classrooms, the study offers an empirically grounded basis for designing more relevant university-based teacher training and support initiatives. The findings are intended to inform ministry stakeholders, university-based PD units, and development partners working to strengthen EFL teacher capacity in Uzbekistan.

2. Literature Review

2.1. EFL Teacher Development in the Uzbek Reform Context

The educational system in Uzbekistan has undergone a significant transformation over the past two decades, with a concerted effort being made to improve the quality of language learning, particularly English (Abdullaev, 2021; Eshchanova et al., 2020; Ibragimov & Muradov, 2022). English language proficiency is viewed as one of the major instruments for facilitating students to compete in the global employment sector and avail themselves of global opportunities. The Uzbekistan government has also made learning the English language a priority, as it has included it in the national curriculum at various levels of schooling (Presidential Decree PQ-4963, 2021; Presidential Decree PQ-5117, 2021; Presidential Decree PF-79, 2023). The Ministry of Public Education has taken steps for improvement in teaching methodologies, textbook revision, and teachers' capacity building to teach English (Education Sector Plan of Uzbekistan, 2019–2023). However, despite these endeavors, there remains a huge gap between policy intentions and classroom realities, particularly in the area of teacher development. According to the Ministry of Education of Uzbekistan (2020), although there have been educational reforms purporting to improve the quality of teaching, there is still an absence of coherence between such reforms and the daily realities faced by EFL teachers in the classroom.

One of the most important determinants of educational quality is teacher professional development, yet a vast majority of EFL teachers in Uzbekistan are faced with serious problems. These problems are manifold and include oversized classrooms, limited access to modern technology, a lack of instructional materials, and a lack of opportunities for specialized

training. In accordance with the British Council report (Dunn & Iwaniec, 2023), the majority of teachers in Uzbekistan still apply traditional methods of teaching, e.g., the grammar-translation method, which prevents them from being able to engage students in communicative and interactive language learning. Moreover, there is a considerable shortage of regular, systematic professional development programs that address the evolving needs of teachers within the framework of a quickly changing educational sector. Although there have been some governmental and international attempts at offering training to teachers, there is no solid evidence that these types of programs are sufficiently tailored to the particular needs of the EFL teachers in the context of Uzbekistan's public schools. A study carried out by Tashkent State University of Pedagogy (2021) informs that professional development courses for teachers are too broad and do not cover specific issues, such as the integration of technology or classroom management, that are key to the success of EFL instruction.

Research in teacher professional development all over the world has focused on needs-based and context-specific training. A research work by Cavendish et al. (2021) concluded that effective professional development programs must be very closely connected to the actual classroom experience and issues of the teachers. This includes offering training in classroom management, student motivation, and integration of technology, all of which are essential in today's world of education (Cavendish et al., 2021). Similarly, the report of the Asian Development Bank (2022) highlighted the need for continuous professional development beyond teachers' initial education and targeted at assisting in the development of teachers' competencies throughout their professional lives. In Uzbekistan, however, it is quite obvious that such programs need to be more specifically designed to address the specific needs of those educators who teach in Uzbekistan's particular educational settings. As the needs of English language teaching evolve, so do professional development opportunities for in-service teachers (Adey, 2004). This study is an effort to fill this gap by revealing Uzbekistan EFL teachers' specific professional development needs, i.e., the challenges they face in developing their teaching practice and students' performance.

2.2. Needs Analysis in EFL Education

Needs analysis is a tradition well established in applied linguistics and language education, a tradition which was developed to ensure that the provision made meets the real needs, constraints, and objectives of the learners or professional groups concerned. The foundational work of Richterich & Chancerel (1978) stressed the systematic identification of needs in language teaching as a basis for more relevant approaches to curriculum and program design. This foundation was extended by Hutchinson & Waters (1987), who proposed a distinction between target needs and learning needs. Target needs, as defined by Hutchinson & Waters, are what the learners or the professional participants concerned need to be able to do in the target situation, while learning needs are what the learners concerned need to be able to learn. The distinction is particularly relevant to the present study, as the needs of EFL teachers as professionals can only be seen as a function of what teachers want to achieve as teachers but also as a function of the circumstances in which teachers learn, practice, and apply what they have learned as teachers.

Subsequent studies have extended the scope of needs analysis beyond the basic deficit approach. West (1994), for instance, has produced a comprehensive typology of needs analysis approaches and has demonstrated that needs can be analyzed using a variety of perspectives, including present situation analysis, target situation analysis, deficiency analysis, strategy

analysis, means analysis, and language audits. Long (2005), for example, has further emphasized that in order to conduct a quality needs analysis, it is important to collect multiple forms of data rather than relying solely on a single instrument based on participants' self-reports. This is because different forms of data can highlight different aspects of need. In this way, needs analysis is not just about asking participants to report their needs; rather, it is a systematic approach to understanding participants' perceived needs, contextual constraints, and task-related demands in order to develop more valid and useful educational interventions (Long, 2005; West, 1994). This understanding of needs analysis is highly relevant to teacher professional development because what teachers report as important needs, what classrooms objectively require, and what training systems can support do not always coincide.

In the field of English language education, needs analysis has had a particularly significant influence on English for Specific Purposes and teacher learning situations. This is because needs analysis can offer a rationale for tailoring language instruction and support to local needs. According to Hyland & Hyland (2006), effective needs analysis involves close attention to the social, institutional, and communicative contexts of language use and language learning. In the field of EFL teacher education, this suggests that teacher education should not be developed as a generic entity separate from the teachers' own working environment. Rather, it should be shaped by the actual instructional needs of the teachers, including factors like class sizes, learner diversity, materials availability, language proficiency, institutional requirements, and technological circumstances. In the field of teacher development, this approach also corresponds with the perspective that teacher professional growth can be most productive when contextual, application-centered, and responsive to teachers' immediate instructional needs (Richards & Farrell, 2005; Farrell & Jacobs, 2020).

This framework is particularly pertinent in Uzbekistan because existing research suggests that teacher PD is still largely top-down in nature and is not sufficiently connected to teachers' real-world realities. Research in Uzbekistan suggests that EFL teachers often report a disconnection between PD content offered and teachers' real-world realities in terms of managing classrooms, using technology, communicative language teaching, and supporting mixed-ability classrooms (Akhadaliyevich & Nozimakhon, 2022; Kosimov, 2021; Kosimov & Latipov, 2022; Tashkent State University of Pedagogy, 2021). More recent research suggests that teachers need PD that is even more sensitive to context, practical in nature, and connected to teachers' daily realities rather than generalized PD approaches that assume all teachers need the same thing (Mamadaminova & Khadjikhanova, 2023; Dunn & Iwaniec, 2023). From the needs analysis perspective, this implies that the development of the PD plan in Uzbekistan should take into account not only the needs of the teachers, as perceived by them, but also the objective conditions of the classrooms, and the pedagogical needs of EFL teaching as an integral part of the needs analysis.

In the present research, the needs analysis was not considered simply as an initial step for administrative purposes, but as the very basis for the development of an appropriate PD plan. The research aims to examine the needs of early career in-service EFL teachers in public schools, as perceived by the teachers, and its relation to the objective conditions of the classrooms, with the ultimate aim of informing an appropriate PD plan. In this manner, this study makes use of the established needs analysis tradition to inform not only the methodological design but also to legitimate the pragmatic goal of developing a more relevant and context-specific model of teacher professional development (Hutchinson & Waters, 1987; Long, 2005; Hyland & Hyland, 2006).

2.3. Existing PD Models, Limitations, and Design Implications

Teacher professional development in Uzbekistan has witnessed significant growth in recent times, as part of the larger efforts aimed at improving the quality of teaching in the country, as well as responding to the changing needs of the nation. There has been an expansion in the opportunities for teachers to access professional development programs, including those initiated by the government, institutions, and even online learning, as acknowledged in the recent report by the Ministry of Education (2022). There are also indications that internationally supported programs have helped introduce modern pedagogies, including communicative language teaching and task-based learning, and that teachers are also being exposed to new teaching practices and professional standards, as acknowledged in the studies by Cavendish et al. (2021) and Dunn & Iwaniec (2023). It is, therefore, an indication that professional development has come to be recognized as an essential aspect in the development of the education sector in the country, as acknowledged in the study by Mamadaminova & Madalińska-Michalak (2025).

However, it has also been indicated in the literature that the effectiveness of the available professional development programs is also hindered by the continued gap between the available professional development programs and the actual classroom situation. According to all the local and international literature, the common pattern observed is that the available professional development programs are too general, too theoretical, and too lacking in terms of support for the actual classroom situation in which EFL teachers are required to work (Nazirova et al., 2023; Mamadaminova & Madalińska-Michalak, 2025). In the Uzbek context, this has been particularly observed in terms of dealing with large classes, technology, classroom support, and teachers' own language confidence (Dunn & Iwaniec, 2023; Tashkent State University of Pedagogy, 2021). Even in those contexts in which teachers are given short-term exposure to new ideas in teaching through international or donor agency programs, it has been observed that the sustainability of these programs is hindered in the actual classroom situation (Cavendish et al., 2021; Sadeghi & Richards, 2021).

Another limitation, in this regard, relates to the general tendency of PD models to follow the 'one-size-fits-all' approach in terms of structure. As indicated by the findings of the studies carried out in Uzbekistan, there are still ample training opportunities that remain too generic in nature, failing to adequately address the daily problems that EFL teachers face in the classrooms, such as dealing with learner diversity, coping with the lack of technological tools, and keeping communicative interactions in the classrooms (Mamadaminova & Khadjikhanova, 2023; Kosimov & Latipov, 2022). This concern may be particularly relevant in the case of early-career teachers, as they may need not only theoretical but also practical support as they start to build their professional routines and identities in the classrooms (Richards & Farrell, 2005; Sadeghi & Richards, 2021).

Collectively, the studies indicate that there are several design implications for PD in Uzbekistan that can lead to greater effectiveness in EFL teaching. Firstly, PD for EFL teachers should be contextualized in the reality of overcrowded classrooms, resource scarcity, and mixed-level classrooms rather than in the idealized teaching situation (Dunn & Iwaniec, 2023; Mukhiddinova & Agzamkhodjaeva, 2025). Secondly, PD programs should go beyond the lecture method and incorporate demonstration, rehearsal, guided practice, and opportunities for immediate application (Richards & Farrell, 2005; Cavendish et al., 2021). Thirdly, support for

transfer is as important as the content of PD, and thus mentoring, observation, and follow-up are necessary for teachers after the PD program (Sadeghi & Richards, 2021; Mamadaminova & Madalińska-Michalak, 2025). Finally, while international frameworks and models are useful for PD, they should be contextualized for Uzbekistan in order for them to be relevant and sustainable in the Uzbek school context (Kosimov, 2021; Tashkent State University of Pedagogy, 2021).

2.4. Research Gap / Conceptual Justification for the Present Study

Overall, the literature examined here indicates that despite the growing policy importance of teacher professional development in Uzbekistan, the level of relevance of the available teacher professional development to the professional needs of EFL teachers, especially young ones, still leaves much to be desired. Indeed, all the literature examined here highlights similar issues such as the need to deal with large mixed-ability classes, technological difficulties, students' speaking skills, and the general impracticality of teacher professional development courses (Mamadaminova & Khadjikhanova, 2023; Dunn & Iwaniec, 2023; Tashkent State University of Pedagogy, 2021). At the same time, the literature examined here indicates that the available teacher professional development models and frameworks may still be too general to deal with these issues in a detailed and sensitive manner (Kosimov, 2021; Cavendish et al., 2021). What seems to be lacking here is the empirical basis for understanding what teachers perceive as their urgent professional development needs and how these needs could be turned into a feasible and sensitive model of teacher professional development.

The other significant gap relates to the lack of a comprehensive needs assessment as an essential element of planning for PD. Several research findings indicate that, in Uzbekistan, professional development planning and provision seem to be top-down and policy-oriented, with insufficient attention given to teachers' actual classroom constraints and perceived priorities (Akhadaliyevich & Nozimakhon, 2022; Kosimov & Latipov, 2022; Mamadaminova & Khadjikhanova, 2023). This implies that professional development provision may not be aligned with actual teachers' needs for enhancing their practice. This is particularly critical for EFL instruction, as it does not rely simply on general pedagogy but also incorporates English language instruction and other elements (Farrell & Jacobs, 2020; Richards & Farrell, 2005).

The current study fills this information gap by carrying out a specific needs analysis among in-service EFL teachers in the initial stages of their teaching career in public schools within the Olmazor district of Tashkent city. Instead of taking the existing models of PD as the best possible ones, the current study examines the lived experiences of the teachers' constraints in the classroom setting, as well as the teachers' perception of the existing PD support as effective or not effective in addressing the teachers' needs, and the nature of support that the teachers find the most relevant to the teaching context. By moving from the level of needs analysis to the level of design implications, the current study can provide a practical contribution to the existing literature on the PD of EFL teachers in public schools in Uzbekistan (Mamadaminova & Madalińska-Michalak, 2025; Sadeghi & Richards, 2021).

2.5. International Collaboration and Training Programs

Global partnership has an important part to play in the professional development of Uzbekistan EFL teachers, positioning global best practice and creative approaches to teaching at the fore. Projects such as the English Speaking Nation: Secondary Teacher Training (ESN: STT) (American Council, 2021-2023), Teaching for Success (British Council, 2017-2023), and the

English for Teachers (British Council, 2017-2019) programs have placed innovative approaches to teaching, such as task-based learning and communicative language teaching, into the teaching practice of Uzbekistan EFL teachers. These international initiatives sponsored by bodies such as the American Councils and the British Council provide opportunities for teachers to expand their pedagogic skills and pedagogy. However, according to Avalos (2021), even though such collaborations have brought immediate benefits, they have little impact on long-term classroom quality because of insufficient sponsorship by successive follow-up training. These foreign programs that are funded by bodies such as the American Councils and the British Council become channels through which teachers' teaching expertise and skills in pedagogy are enhanced (ESN-Teachers, 2023). Even though, as averred by Cavendish et al. (2021), such partnerships are making short-term achievements, they do little to impact long-term teaching quality since they are not funded by ongoing follow-up training (Cavendish et al., 2021). The majority of teacher's report that while they acquire new techniques through these courses, they do not properly implement them once they return to their classrooms, mainly due to the fact that training is poorly adjusted to their own particular local contexts of teaching.

Furthermore, foreign training courses tend to focus primarily on language abilities and practice in the classroom but overlook the larger issues Uzbek teachers must contend with, such as managing gigantic classes or adapting to the lack of teaching materials. Based on a British Council (2023) report, while these courses have introduced positive pedagogies, these tend to be disconnected from Uzbek classroom realities. The inability to implement theoretically what is acquired in programs in real teaching settings means that the effective work of such programs cannot be continued in the long term. The remedy for this deficiency is to have more context-based models of training that take into account the particular limitations of EFL teachers of Uzbekistan. This would include continuous support through mentorship, classroom visits, and peer working to ensure that the teachers are effectively able to adopt new practices into their classrooms.

3. Aims, Purpose and Objectives

This exploratory-descriptive study aims to generate an evidence-based foundation for a context-sensitive PD framework for early-career EFL teachers in Uzbekistan. Specifically, it pursues the following objectives:

- Objective 1 (Context profiling): Describe the eligible teacher population's key background characteristics relevant to PD planning. Data source/method: Context-demographics questionnaire.
- Objective 2 (Needs identification): Identify teachers' most urgent PD needs and constraints. Data source/method: FGDs (5 groups of 5) analyzed via thematic analysis.
- Objective 3 (Perceived gaps in current PD): Examine teachers' perceptions of usefulness and gaps in existing PD provision. Data source/method: FGDs (same dataset; thematic analysis).
- Objective 4 (Design implications): Translate identified needs into a feasible PD prototype (modules, delivery modes, monitoring indicators). Data source/method: Synthesis of Objectives 1–3 into a PD prototype table (Appendix G).

4. Methodology

The approach taken in this study was exploratory-descriptive, characterized by a focus on ensuring the study is qualitatively oriented. A short context-demographics questionnaire was employed for the purpose of briefly characterizing the eligible participants and finding background information to situate the findings from the focus group. The dominant source for data, however, came from the conduct of a series of semi-structured focus group discussions (FGDs) to explore perceived professional development needs, constraints, and preferences.

4.1. Sampling and Participants

Olmazor District currently has around 340 in-service EFL teachers. To make the results of the analysis of learning needs useful for formulating a relevant development program for in-service EFL teachers in this research study, a series of eligibility criteria has been employed in gathering a relevant sampling of EFL teachers in Olmazor District. The two conditions for being eligible for this study are:

- (1) The teacher possesses a B2 or higher English proficiency certification.
- (2) The EFL teacher has no more than five years of experience in EFL instruction.

Based on the official list provided under the Olmazor District Department of Education, a total of 72 EFL in-service teachers were eligible participants; 25 teachers were randomly selected using a fishbowl method to participate in the study. These 25 teachers formed the final study pool ($n=25$) and were divided into five focus group discussions, with five participants per discussion. All instruments of data collection, such as questionnaires and focus group discussions, were only carried out with this pool of 25 participants.

Table 1. Recruitment Flow Table

Stage	n
Estimated EFL teacher population in Olmazor District	340
Eligible teachers meeting inclusion criteria (B2+; ≤ 5 years' experience)	72
Randomly selected via fishbowl method for participation	25
Completed demographics/context questionnaire	25
Participated in FGDs (5 groups $\times$ 5)	25

Because participants were selected randomly from the eligible list ($N = 72$), the sample reduces selection bias within the defined early-career B2+ subgroup, although findings are not intended to represent all EFL teachers in Olmazor District.

4.2. Instruments and Data Collection

At the needs analysis stage, two significant tools were employed: a questionnaire to collect demographic data and contextual data from the participants and semi-structured interviews with a smaller segment of the participants. This was done to hold FGDs to explore their professional development needs in greater depth. Questionnaire: The questionnaire was used as the first instrument in a sample of 25 teachers from public schools, with each holding a B2+ language certificate and a maximum of five years of teaching experience. The questionnaire

provided baseline demographic data and contextual information regarding education, experience as a teacher, classroom structure, access to instructional tools, and prior involvement in professional development activities. This tool gave important information on the teachers' credentials and work environment, which are important determinants of their professional development requirements.

Focus Group Discussions: FGD was used to identify the professional needs of in-service EFL teachers. The researcher stratified the 25 participants into five groups with five members each. The researcher interviewed each of the groups separately in June 2025 at a school that was convenient for each FGD participant, and their responses were recorded using a smartphone recorder to ensure accuracy for future thematic analysis. The objective was to gather different insights and opinions regarding the perceived needs and training issues/challenges of the in-service EFL teachers. In these FGDs, a researcher acted as the interviewer/moderator, assisted by a guide consisting of semi-structured format questions. These FGDs took approximately 30-45 minutes to conduct. A brief introduction about the study and the procedures involved preceded the actual study for all the teachers before they provided their consent. The rules were established from the start to promote teacher honesty in the study. These guidelines included the fact that the discussion would be voluntary, the responding should be respectful, and the participant could always choose to bypass the discussion at any point. It was stressed that there were no right or wrong answers and that the data would be collected and presented anonymously (using group codes or pseudonyms instead of names or institutions). The recordings of the audio files were all saved to facilitate research.

As FGDs were conducted in English, no translation/back-translation was required. All sessions were audio-recorded (smartphone recorder) and transcribed for analysis. Each of the participants had a B2+ level of English-speaking proficiency, which aided in their effective presentation of their personal opinions in English. The transcripts of these discussions were typed by the researcher using an intelligent verbatim transcription style, whereby the actual meanings of the statements made by the participants remained intact, omitting filler statements and repetitive statements with unaltered meanings. Member checking. To further enhance the credibility of our work, we conducted member checking. The participants were asked to review the quotes extracted (and the short summary interpretations we use in the reporting) for accuracy and to check that no intended meaning has been lost. Appendix F displays a sample documentation of the member-check process from one participant.

4.3. Data Analysis

The data from the focus group discussions was analyzed using reflexive thematic analysis in six steps outlined by Braun and Clarke: becoming familiar through repeated listening and reading of the data, generating initial codes, considering potential themes, reviewing the themes in terms of the coded extracts and the data set as a whole, defining and labeling the themes, and finally preparing the report with quotes from the data. The method was exploratory and descriptive in essence to determine teachers' professional development needs, what shortcomings exist in the professional development services that teachers are exposed to currently, and how the design of a professional development framework would be informed.

Coding was done through three intensive cycles: open → axial → selective coding. In Cycle 1—open coding—codes were applied to each segment of the transcripts to identify meaningful units representing challenges, limitations, and preferences of teachers for PD. In Cycle 2—

axial coding, codes were grouped into categories according to their underlying dimensions (mentor teacher support categories), like transfer barriers, facilitators, and PD preferences. In Cycle 3—selective/development coding of themes—categories were developed further to produce the final themes to organize the findings of this study and to provide a basis for the structure of the PD framework proposal. A codebook was developed through cycles to keep track of code labels/characteristics, code definitions, criteria for inclusion & exclusion, and exemplar quotes.

The handling of the coding and text excerpt searches was done using the software program, Quirkos. Microsoft Excel software was used to track changes in the codebook and to create a summary matrix of information for transparent reporting. The transcripts underwent coding by the principal researcher. In the absence of multiple coders, percent agreement or Cohen's Kappa statistics on inter-coder reliability could not be calculated. To enhance reliability and ensure transparency, the construction of the codebook is documented in an appendix (Appendix E), and an audit trail of research decisions is maintained. This process is further supported by the validation method of 'member checking' the results through an appendix submission (Appendix F). Additionally, 'triangulation' of the coded results is achieved by incorporating backdrop information from the candidate population obtained through a questionnaire.

4.4. Trustworthiness

There were five focus group discussion sessions conducted, consisting of five participants in each session. This was done in order to ensure manageable discussion and gather as many perspectives as possible from the pool of interested participants. Our target for the analysis was 25 participants. This was because it seemed the largest viable number for the impending PD program and represented a fair representation of the district-level volunteers. Informal assessment of code saturation in the analysis stages indicated when new incident codes began ceasing to emerge in the groups. By the fourth and fifth FGDs, there was no landmark of new incident code generation, and all of the remaining sessions were for filling gaps in the theme generation. Hence, five FGDs were sufficient for our analysis.

Multiple strategies were employed to enhance the trust and trustworthiness of the results. The first method used was member checking; this entailed encouraging the participants to validate the collected quotations (and associated shortened explanation interpretations as used in the reporting) for the sake of authenticity and ensuring the correct message had been conveyed; see Appendix F. Triangulation entailed utilizing the contextual elements provided in the questionnaires by the eligible population to validate the results of the focus group themes using the concerns in the background. The audit trail was maintained throughout the analysis in order for it to meet the criteria of dependability and confirmability. This involved maintaining dated copies of the development of the codebook, analysis memos, and alterations related to themes/code revisions (managed in Microsoft Quirkos & Microsoft Excel).

Figure E of the appendix E contains a sampling of what the completed codebook contains (coding titles, definitions, and cut-offs for coding themes, including exemplar quotes). Peer debriefing was not formally done, since the analysis process was carried out by only one researcher, without auditing by an external qualitative analyst. This is counted as one of its limitations. As a follow-through, whether in resubmitting or in pursuing other lines, peer debriefing or auditing by an external party for theme definition and their supporting texts will be done.

Researcher Reflexivity and Positionality. FGD discussions, as well as the analysis, were facilitated by the researcher (affiliated with Tashkent Metropolitan University). She is trained in TEFL/TESOL and pursuing her doctorate in Curriculum and Instruction, thereby providing expertise in leading English-medium teaching and learning, as well as in understanding professional development needs. To ensure bias was not present, she emphasized volunteer participation, protected confidentiality, and made sure no one's employment or performance evaluations were impacted by their participation. To analyze the data, she maintained reflective logs of all assumptions she made, decisions she undertook, and readings she made, and these were all linked to direct quotes she obtained from respondents. Member checking was conducted for these quotes and the resultant overall findings in order to ensure the original meaning was captured.

4.5. Ethical Considerations

The study targeted adult in-service teachers, investigated professional development needs, and presented minimal risk. Each participant of the study signed a written consent form (Appendix A) prior to data collection. Participation was completely voluntary, with clear assurances that they might prefer not to answer a question or withdraw at any time without penalty. The researcher established no abuse of power; there was respectful dialogue where there would be no judgment as to right or wrong answers, and shared experiences were to remain confidential. All information transcribed and reported was anonymized: participants obtained codes that were non-identifying; school names were not provided; and when quotes were used, all identifying details had been removed. Audio files and transcripts were stored securely and used only for research purposes.

Ethical review was provided by an exemption/waiver of ethics from Tashkent Metropolitan University management and confirmations in writing that the research project presented a minimal risk and would be performed under usual precautions: informed consent, confidentiality, anonymization, and data protection. The signed Exemption/Waiver Letter is attached in Appendix B. Data availability. Audio files and transcripts will not be shared or deposited due to confidentiality and the risk of possible identification. The excerpts in the manuscript and the respective codebook excerpt will only be deposited in the appendices.

5. Findings

5.1. Demographic Data and Contextual Information

Breakdown by Age: Most English language teachers who participated in the study were relatively young. Approximately 64.0% were between 22-30 years old, which made up the highest number, represented by $n = 16$. The group that was between 31-35 years old represented 28.0% ($n = 7$), while those that were between 36-40 years old comprised 8.0% ($n = 2$). The sample represented in the study reflects early-career teachers, offering a wide age range comprising young and more senior teachers to ground the study regardless of the stages involved.

Language Ability Certificates: The questionnaire also inquired about the teachers' certificates of language ability. The one that most teachers named was CEFR-based, since 41.4% of the

teachers referred to this certificate. IELTS was indicated by 24.1% of the respondents, and 13.8% referred to APTIS. Interestingly, some teachers also added that they have taken other language ability tests, such as Duolingo or local ones, and each of these represented 13.8%. A minority of respondents reported never having taken a formal language proficiency exam. This finding is evidence that there is a varied set of English certification backgrounds among participants and that while a large majority might be accustomed to standardized language test structures, there are others who have minimal formal records to verify language abilities. These findings can have design implications for designing differentiated professional development resources that are able to respond to variation in the language proficiency of teachers.

Teaching Experience: All participants in the focus group belong to the early-career group, with 1 to 5 years of working experience. Based on Table 2, almost equal numbers belong to the 3-year experience level ($n = 12$, 48%). The 4-year level was comprised of 5 participants (20%), while 3 participants each belong to the 1-year (12%) and 2-year (12%) experience levels. Another smaller sector reported 5 years of working experience ($n = 2$, 8%). Overall, this data indicates the early-career professionals who took part in this research study remain relatively new to their crafts, which include establishing the day-to-day methodologies in their line of work.

Table 2. Teaching experience of FGD participants ($n = 25$)

Years of teaching experience	n	%
1 year	3	12
2 years	3	12
3 years	12	48
4 years	5	20
5 years	2	8
Total	25	100

5.2. Focus Group Discussion Results

5.2.1. Findings of the Focus Group Discussions (FGD)

The Focus Group Discussions (FGD) with 25 in-service EFL teachers yielded valuable insights into the requirements of their professional development, which were subsequently used to inform decisions regarding the design of the professional development (PD) program. The thematic content analysis of the FGD responses also revealed some salient recurring themes that indicate the key issues and concerns of the teachers. The principal themes invoked, with reflective quotes extracted from the FGDs, are listed below.

5.2.1.1. Large Mixed-Ability Classes

The overall problem at the FGDs was managing huge classes with varying levels of ability of the students in them. Teachers seemed to worry about whether they would manage to provide varying students with individualized attention in one class, particularly individualized care and discipline within the classroom.

Sample excerpt:

"It is difficult to make all of them participative if you have 35 students in a class." You have students who are advanced, and others do not even know how to speak English. I do not know how to manage differences this way effectively."

– Teacher 3, FGD Group 1

This finding affirms previous research conducted by Mamadaminova & Khadjikhanova (2023), wherein the researchers proved that large, mixed-ability classes are one of the most common challenges facing EFL teachers in Uzbekistan. The FGD members said that they did not possess ways of teaching all students at the same time, which is also one of the essential requirements of professional development.

5.2.1.2. Technology in Teaching

There was also a prevalent theme that we could find in the FGDs, i.e., no use of technology in the classroom. Though many of the instructors wished to use online materials and technology tools, they complained of not having access to technological facilities and of receiving training in how to use technology well in teaching.

Sample Excerpt:

"I would like to use more web resources, but we just do not have the equipment. Even if we had computers, I would not know how to use them in a proper way while teaching."

– Teacher 7, FGD Group 2

The experts referred to a lack of digital materials and professional education to utilize them efficiently as an issue. This statement validates the need for PD courses that provide both pedagogical and technological support to incorporate digital material into language teaching. Shaturaev (2021) also found similar limitations in his study on Uzbekistan EFL teachers.

5.2.1.3 Speaking Confidence and Communicative Practice

The majority of teachers were worried about their own proficiency in the language, specifically their speaking ability. This was felt to be personally impacting their ability to teach communicative language skills well. The teachers also pointed out that they employed a lot of non-communicative teaching methods, such as the grammar-translation method, since they were not sure of adopting more communicative and interactive methods of teaching.

Sample Excerpt:

"I am confident in grammar, but if I have to speak English fluently in front of students, I feel nervous. It's hard to teach speaking if you're not confident yourself."

– Teacher 10, FGD Group 3

FGDs indicated that building speaking confidence should be an area of emphasis for PD programs since the majority of teachers are not confident when communicating orally in English with their students. This aligns with the focus of Tashkent State University of Pedagogy (2021), which underscored building teachers' language capabilities to improve their teaching effectiveness.

5.2.1.4. Practical, Context-Specific PD

One of the most frequent themes throughout the FGDs was that teachers indicated that they wanted PD courses that are not only theoretical but also immensely practical and applicable in their very own classroom realities. Teachers required professional development that would provide them with concrete techniques that they could apply immediately in their very own classrooms.

Sample Excerpt:

"I've attended numerous workshops, but they don't really instruct me on how to manage my classroom or integrate technology in practice. I need training that is more realistic and context-specific in our situation."

– Teacher 12, FGD Group 4

This criticism underscores the need for the creation of PD programs directly from the day-to-day life of teachers in the classrooms. Teachers, according to Shavkatova & Mirzaahmedov (2025), believe that PD programs are not practical enough and are not relevant to their daily classroom issues.

5.2.1.5. Lack of Support for Continuous Professional Development

The FGDs also suggested that the majority of the teachers feel that they do not receive sustained support after professional development activities. Teachers called for ongoing mentoring and follow-up support to facilitate their applying what they have learned in training to teaching practice.

Sample Excerpt

"We attend these workshops, yet when they're over, there's no follow-up. You can't do everything without someone checking on you or reminding you where you stand."

– Teacher 15, FGD Group 5

This finding calls for PD programs offering long-term support, for instance, in the form of mentoring or peer support, which can help teachers reinforce and implement the skills and information acquired during initial training. A lack of this type of long-term support is a common issue for PD programs globally (Cavendish et al., 2021).

Results of the FGD illuminate what professional growth EFL teachers in Uzbekistan require. Classroom management, integration of technology, and mastery of languages are concerns for teachers, which impact their success as instructors. Further, there is a high demand for PD

programs that are contextual, applied, and offer sustained support. These findings hold tremendous implications for designing a professional development program that can effectively target such matters and assist teachers in their professional growth. The next phase of this study will be to use these findings to inform the design of a targeted PD program for improving the quality of English language instruction and teaching practices in Uzbekistan.

Table 3. Theme suggestions for the PD development

FGD Theme	Key need expressed by teachers	Implications for PD design
Large, mixed-ability classes	Differentiation and discipline strategies feasible in large classes	Low-prep differentiation routines, mixed-level task design, behavior-support systems, rehearsals via micro-teaching
Technology constraints	Lack of equipment and limited know-how	“Minimum viable tech” using phones/offline tools; no-tech alternatives; ready-to-use digital/printable templates
Speaking confidence & communicative practice	Anxiety about speaking + reliance on grammar-translation	Classroom English bank; coaching and micro-teaching; speaking-task facilitation practice; feedback cycles
Practical, context-specific PD	PD perceived as too theoretical	Demo lessons, rehearsal, problem-based tasks, materials adapted to local textbooks, and immediate classroom transfer focus
Lack of follow-up support	No sustained post-training support	PLC circles; peer observation cycles; mentoring/coaching check-ins; structured follow-up tasks

6. A Need-Based PD Framework Prototype

Building on the findings of the context-demographics questionnaire and the five themes identified through the FGDs, this study developed a needs-based professional development framework prototype for early-career CEFR B2+ EFL teachers in Olmazor District. The prototype was designed as a practical response to the most consistently reported challenges in participants’ professional lives: managing large mixed-ability classes, teaching under technological constraints, low confidence in speaking and facilitating communicative practice, dissatisfaction with theoretical and insufficiently contextualized PD, and the absence of sustained post-training support. Rather than proposing a generic training package, the framework translates identified needs into five focused modules with corresponding learning outcomes, delivery modes, and monitoring indicators. In this sense, the prototype functions as the study’s applied output, showing how a qualitative needs analysis can be converted into a feasible and context-sensitive teacher development design (Richards & Farrell, 2005; Sadeghi & Richards, 2021).

6.1. Managing Large, Mixed-Ability Classes

The first module deals with the challenge of teaching in large classes with significant variations in students' language proficiency, which was one of the most prominent issues that emerged during the FGDs. In this regard, the participants reported that it is hard to ensure students' active engagement, differentiate instruction in such large and heterogeneous classes, and discipline students. This is in line with previous research that indicated that large and heterogeneous classes are one of the most persistent issues faced by EFL teachers in Uzbekistan (Mamadaminova & Khadjikhanova, 2023; Mukhiddinova & Agzamkhodjaeva, 2025). Therefore, the module on teaching in large classes with significant variations in students' language proficiency is designed with low preparation differentiation routines, flexible groupings, scaffolding, and classroom management routines that are feasible in large classes. In addition, the module is designed with a focus on demonstration lessons, guided deconstruction of teaching actions, microteaching, and peer feedback, followed by the generation of lesson plans that are ready for the classroom. This is in line with the general argument in PD research that emphasizes that teachers benefit most from learning opportunities that are closely linked with their immediate classroom realities (Richards & Farrell, 2005; Cavendish et al., 2021).

6.2. Low/No-Tech Teaching

The second module is a direct response to participants' interest in technology, as well as their repeated reports of poor infrastructure and insufficient pedagogical support in technology tool utilization. Indeed, previous studies on Uzbek EFL teachers have pointed out that these teachers often work in environments characterized by poor access to modern technological tools, as well as insufficient support in technology tool utilization for teaching EFL (Dunn & Iwaniec, 2023; Shaturaev, 2022). The results of the current study indicate that a realistic PD program response to participants' concerns should not be premised on the assumption of teachers' access to high-tech tools, but rather on teachers' productivity in such environments. Thus, the second module in the proposed PD program has been developed on a minimum viable tech philosophy, aiming to train teachers in adapting textbook materials into interactive no-tech and low-tech activities, as well as in utilizing realistic tools such as telephone audio, offline tools, and printable templates, for example. This is a direct response to participants' call for strategies that take into account teachers' and schools' limitations, as well as those that expand pedagogical possibilities (Farrell & Jacobs, 2020; Sadeghi & Richards, 2021).

6.3. Teaching Speaking and Teacher Speaking Confidence

The third module is based on the premise that teachers own speaking anxiety and limited classroom English confidence are factors that influence their ability to implement communicative learning environments. Participants were asked to report their nervous feelings about speaking English in front of students and their tendency to overuse grammatical approaches in their teaching. This is also supported by previous research that indicated limited speaking confidence in teachers may influence their ability to model language use and implement communicative approaches (Tashkent State University of Pedagogy, 2021; Farrell & Jacobs, 2020). In response to this, the module is designed with two main purposes: increasing teachers' confidence in using classroom English and increasing their ability to implement structured speaking activities such as pair work, information gap activities, and role-plays. The proposed model is based on coached microteaching approaches and scripted classroom English. The rationale behind this module is that communicative approaches cannot be implemented without increasing teachers' confidence in using classroom English and

providing them with opportunities to practice and develop routine language use in their instruction (Richards & Farrell, 2005; Sadeghi & Richards, 2021).

6.4. Teaching Speaking and Teacher Speaking Confidence

The third module is informed by the idea that teachers' own speaking anxieties and low levels of English confidence in the classroom influence their capacity to promote communicative learning environments. The participants expressed concerns about speaking English in front of learners and suggested that this may have contributed to a reliance on grammar-focused practice over more speaking-focused practice. This is similar to previous studies that suggested teachers' low levels of speaking confidence can influence teachers' capacity to promote speaking and communication in the classroom (Tashkent State University of Pedagogy, 2021; Farrell & Jacobs, 2020). As such, this module addresses two related needs: teacher confidence in using English in the classroom and teacher competence in supporting structured speaking activities such as pair work, information gap tasks, and role-plays. The intended mode of delivery is coached microteaching cycles, scripting of English in the classroom, rehearsing, and providing feedback. The rationale of this module is that communicative approaches in language teaching cannot be divorced from teacher confidence; teachers are likely to adopt methods that emphasize speaking when they are given opportunities to rehearse these methods and develop routine language for teaching (Richards & Farrell, 2005; Sadeghi & Richards, 2021).

6.5. Communicative Methods Under Constraints

The fourth module deals with the general concern that much of what is offered in terms of professional development is considered too theoretical and not well adapted to the realities of the local classrooms. The participants emphasized the need for professional development that is realistic, applicable, and immediately useful in relation to the problems of the classroom. This is also in line with the research on PD in Uzbekistan, where the experience was considered too general and not well adapted to the actual conditions of the teachers (Mamadaminova & Madalińska-Michalak, 2025; Kosimov & Latipov, 2022). The module will deal with the redesign of grammar- or textbook-bound lessons into more communicative sequences of lessons, keeping in view the actual constraints of the classrooms, such as large numbers of students, mixed proficiency, and low student motivation. The rationale for the module is based on the understanding that teachers do not only need to be exposed to modern methodologies such as communicative language teaching, but they also need support in how to apply these methodologies in the classroom, keeping in view the actual conditions of the schools (Farrell & Jacobs, 2020; Cavendish et al., 2021).

6.6. Transfer Support: Follow-Up and Community of Practice

The fifth module addresses the recurring problem of professional development workshops ending with the workshop itself, without any further support or encouragement for teachers to continue with implementation. Participants indicated that they need mentoring support, as well as opportunities to discuss problems faced after attending the workshop. This is another recurring problem in professional development, as indicated in other literature as well. The PD effect is considered to be reduced if teachers go back to school without reinforcement, coaching, or collaborative follow-up structures in place (Cavendish et al., 2021; Mamadaminova & Madalińska-Michalak, 2025). In order to resolve this issue that has

continued to plague teachers during their professional development, the final module discusses the need to offer support during the transfer of what has been learned by providing teachers with peer observations, PLC circles, discussions for reflection, and problem clinics. This is especially important since the findings of this study have shown that even effective training may not be sustained if teachers return to school without feedback, reinforcement, and support structures (Sadeghi & Richards, 2021; Richards & Farrell, 2005).

6.7. Feasibility

This prototype should be seen as a context-sensitive design proposal informed by the FGD findings from early career CEFR B2+ EFL teachers in a single urban district of Tashkent, as opposed to a final product for national-level adaptation. The minimum requirements for the framework to be implemented in the future might include a trained facilitator for a manageable number of teachers, access to a space for demonstration lessons and micro-teaching, some basic printable materials, teachers' own phones for some activities, and a short-term follow-up phase like a three-month PLC or peer observation schedule. The design of the prototype was intended to be resource-sensitive enough to be implemented in low-resource contexts by prioritizing low-prep approaches, no-tech or low-tech solutions, and school-based approaches over infrastructure needs. However, for the framework to be adapted for use in other, possibly more dissimilar school contexts, including those in the countryside, some level of adaptation might be required (Dunn & Iwaniec, 2023; Mamadaminova & Madalińska-Michalak, 2025).

7. Discussion

The present study aimed to reveal the professional development needs of in-service early-career EFL teachers in public schools in Olmazor District and to transform these needs into a feasible professional development prototype. The results indicate that the needs of the teachers are not limited to the development of specific skills but rather indicate a broader misfit between the current professional development opportunities and the actual context in which these teachers are teaching. In all the themes, the teachers' professional reality has been characterized by the teaching of large classes with mixed-ability students, the variable access to technology in the classroom, the low level of confidence in using English in the classroom, the teachers' dissatisfaction with the overly theoretical nature of the training courses, and the lack of follow-up support after the courses. Collectively, these findings imply that PD for this group should be conceptualized not merely as the delivery of content but as the promotion of adaptation to challenging classroom contexts. This conception of PD aligns with literature that emphasizes the need for teacher professional development to be "practice-oriented, context-responsive, and connected to teachers' daily instructional challenges" (Richards & Farrell, 2005; Cavendish et al., 2021; Sadeghi & Richards, 2021).

A significant contribution of this study is that it does not merely describe the problem but begins to illuminate the logic of what constitutes an appropriate PD response. That is, the findings do not merely imply that teachers need "more training." Rather, they imply that teachers need "different training." For example, the problem of managing large mixed-ability classes is not merely a classroom management issue in a behavioural sense. It also points to the challenge of differentiation, sustaining engagement, and maintaining communicative interactions in overcrowded classes with differing levels of proficiency among the students. This, in turn, helps to explain why the generic workshops might be seen as inadequate, as the teachers are not simply in need of a better level of theory but also a more concrete level of

practice that can be sustained in the face of overcrowding, resource scarcity, and time constraints. This interpretation of the results is consistent with earlier research indicating that large class pedagogy requires rehearsal-based support, as opposed to abstract methodological advice (Mukhiddinova & Agzamkhodjaeva, 2025; Richards & Farrell, 2005). The emphasis on demonstration lessons, low-preparation differentiation, and microteaching, therefore, is a design intervention directly responsive to the form of the identified need, as expressed in the data.

A similar phenomenon can be observed in terms of technology as well. In this case, the participants were not seen to be dismissing the idea of digital pedagogy. Rather, they showed interest in using technology for teaching but also mentioned their current environment in terms of technology. This implies that the main problem here is not the lack of technology. Rather, it is the lack of alignment between the ideal scenario of using technology for teaching and the actual scenario in terms of the availability of resources. This is also the reason for the prototype to be based on the idea of minimum viable tech and not high tech. This is because the main purpose of providing PD in terms of technology should be to enable teachers to make the best use of the available resources, including the lack of resources. This is also in line with the studies that have suggested that the main issue in terms of using technology for EFL settings is not the lack of technology but the ability of teachers to make the best use of the available resources (Kosimov, 2021; Dunn & Iwaniec, 2023; Farrell & Jacobs, 2020).

The results regarding speaking confidence are especially noteworthy, however, since all participants qualified for the study based on this language-related criterion, yet some still evidenced nervousness about speaking English in front of students. Clearly, therefore, formal qualification is not necessarily sufficient to guarantee communicative security in practice. What is important, it seems, is not only language knowledge, as qualified and certified through formal processes, but also communicative confidence in using English publicly and spontaneously as a medium for classroom interaction. In this way, therefore, the problem is not only pedagogical but also professional, since teachers may be less likely to promote communicative activities and more likely to rely on grammar instruction and teacher-centered methods if they lack confidence in their own classroom English. Accordingly, the problem is not simply one of proficiency but also one related to the interrelated concepts of teacher confidence, classroom performance, and communicative pedagogy. Hence, again, the importance of prototype features such as classroom English scripting, rehearsal, micro-teaching, and feedback. The idea here is to support not only linguistic competence per se, but also communicative confidence per se (Tashkent State University of Pedagogy, 2021; Farrell & Jacobs, 2020; Richards & Farrell, 2005).

The theme of practical, context-specific PD serves to further reinforce this interpretation. This is because the participants continually expressed their perception of the current PD opportunities as being too theoretical and not applicable enough to the problems they are facing in the schools. This should not be taken to mean that theory is not important, but rather that theory is important to these teachers in ways that are not yet apparent in the current opportunities. In the Uzbek context, this theme could also be exacerbated by the current prevalence of broad, standardized, or top-down approaches to PD that do not adequately differentiate between subjects, schools, or stages of the teaching career (Mamadaminova & Madalińska-Michalak, 2025; Shavkatova & Mirzaahmedov, 2025). This is to suggest that the quality of PD should not be seen in the volume of content delivered, but rather in the degree to which the teacher can see, try, and apply what is being taught. This is why the prototype has

focused on model lessons, problem-based activities, lesson redesign, and materials that are aligned with local textbooks. This is in direct alignment with the finding that teachers value this over breadth of concepts.

The final theme, lack of follow-up support, is theoretically relevant as it changes the focus from training programs to the transfer of training. The implication of the findings is that the major problem with the current PD system is the lack of support after the training program. The findings of the current study, however, indicate that the problem is actually in the implementation, especially for early career teachers who are still learning to adapt to the demands of the classroom. Therefore, the fact that the prototype included peer observation, PLC circles, coaching check-ins, and tasks is central to the design, as it addresses the social dimension of teacher learning by recognizing the fact that teachers are more likely to change over time with reinforcement, feedback, and support (Cavendish et al., 2021; Kosimov, 2021; Sadeghi & Richards, 2021).

If all five themes are taken into consideration, a coherent picture of the proposed prototype emerges. The framework should not be seen as a supplementary entity added to the results; instead, it can be viewed as a synthesis of the study's findings. The modules on mixed-ability teaching, minimum viable technology, speaking confidence and communicative practice, practical lesson redesign, and follow-up support can be seen as all stemming from a single underlying concept: that any professional development for early career EFL teachers in this setting must be based on what is viable under constraint. This concept can be seen as entirely consistent with the needs analysis approach that underpins the study, where teachers' perceived needs, classroom realities, and implementation realities are seen as interconnected rather than separate domains. In this way, the prototype can be seen as a context-sensitive design proposal based on identified need, rather than as a universal prescription for all schools or all EFL teachers.

It should be noted, however, that this discussion should be contextualized within the boundaries of the study. This is because the participants of this study are early career CEFR B2+ EFL teachers based on one urban district of Tashkent. Therefore, the results are more relevant to this group of teachers than the entire teaching staff of the nation. This is important because, although urban schools based in the capital may be different from those based in other areas, especially rural and remote schools, the results should be used but adapted before application. This qualification of the study is important to ensure that its value is maintained without overgeneralizing the results.

The overall implication of the discussion, therefore, seems to be that the real question is not whether PD is available, but whether it is usable, practice-related, and sustained enough to influence practice in real-life contexts. By linking the identified needs to a prototype that focuses on rehearsal, practice-relatedness, resource-constraint pedagogy, communicative confidence, and post-training support, the study makes a specific but substantial contribution to the design logic of what context-sensitive PD might look like for young EFL teachers in Uzbekistan.

8. Implications and Recommendations

8.1. Implications

The implications of the findings of this study for the design of professional development (PD) for early career EFL teachers in public school settings such as those in Olmazor District, Tashkent, are significant. The limitations of the study, in that it was conducted in only one district of the capital city, mean that the implications of the findings should be thought of in relation to settings such as this, rather than assuming they can be generalized to all settings in Uzbekistan. Second, the continuous shortage of practice PD content is an indication of the need for training modules that can bridge theory to practice in the classroom. Trainers and developers must prioritize experiential learning approaches such as lesson simulation, peer teaching, and case-based problem-solving that enhance pedagogical creativity and self-efficacy.

Third, the technology and access skill deficit reveals a double issue: infrastructural inequality and digital illiteracy. The PD planners should consider including digital skill training in every element as an integral part, not an afterthought. That would make the teachers prepared to utilize the available technology under even restrictive conditions. Fourth, the need for methodological assistance implies a gap between current classroom practices among teachers and contemporary communicative and learner-centered methodologies. PD initiatives in similar contexts should therefore offer prolonged mentoring and follow-up workshops where teachers can internalize and experiment with methodologies such as TBLT, CLT, and differentiated instruction in their respective classrooms.

Fifth, assessment and feedback challenges identify the necessity to revisit teacher assessment literacy. Incorporating applied modules on formative assessment, providing feedback, and performance-based assessment can prepare teachers to more effectively track pupil progress and decrease reliance on summative, exam-based practice. Sixth, the confidence and motivation issues evidenced show that PD programs not only need to augment teachers' skills and knowledge but also their professional identity and confidence. Reflective assignments, systems of recognition, and networks of peers operating through collaboration could be a component of more lasting motivation and a sense of professional belonging.

8.2. Recommendations

On the basis of these implications, some recommendations can be proposed for those concerned with the improvement of professional development of EFL educators in similar contexts to that investigated in this study. These recommendations are based on data collected from CEFR B2+ public school EFL educators in their early career in Olmazor District, Tashkent, and therefore should be considered context-specific, not applicable in other regions of Uzbekistan without further modification.

It is also recommended that mentorship and peer-coaching systems are included in all professional development schemes. Ongoing guidance by experienced teachers and academic mentors would allow teachers to change over time to new pedagogic approaches and adapt them for their own classrooms. This mentoring element is imperative in making short-term training a long-term professional development. Besides, PD sessions must incorporate digital pedagogy training as a core and not an addendum. Enhancing teachers' digital skills and literacy

in using technology in teaching will promote more balanced access to quality education and enhance the technology gap between resourced and under-resourced schools.

The second most important recommendation involves literacy in assessment. Continuing professional development material needs to provide definite potential for teachers to craft formative assessments, design rubrics, and provide constructive feedback in harmony with communicative teaching goals. Acquiring this skill would also allow teachers to monitor learning accomplishments more effectively and move beyond exam-driven assessment practice. Professional development should also address, in particular, teacher motivation and incentives. The implementation of certification levels, micro-credentials, or local award systems could serve as an incentive and keep the subject of continuous learning of interest. Recognition not only motivates based on effort but also generates the feeling of belonging to a professional community of practice.

Greater collaboration between schools and universities is also essential. Greater institutional interrelation can ensure PD material remains connected to ongoing research while being responsive to schoolteachers' lived practice. Collaborative arrangements between schools and universities can also illustrate the co-production of training resources and the exchange of knowledge and expertise among academic researchers and teachers in classrooms. Although these implications and recommendations may provide valuable advice for PD planning in Uzbekistan in general, it is crucial that they are used cautiously beyond the specific context in which they are applied. This is due to the fact that schools in other areas, especially rural areas, may be different in terms of infrastructure, teachers available, class sizes, and access to technology and mentoring support. Therefore, it is imperative that the PD prototype be recognized as a contextual model that requires adaptation and, preferably, piloting in specific contexts before it is adopted at national levels.

9. Conclusion

This study provides an in-depth investigation of the needs of in-service EFL teachers' professional development in Uzbekistan. The findings indicate some of the fields where the teachers require support, and these include classroom management in very large, mixed-ability classes; application of technology in instruction; language skills, particularly confidence in speaking; and having the chance to access more practice-based, context-focused PD courses. The teachers also desired ongoing support and follow-up after PD courses, thus indicating the necessity for continuous professional development.

These findings have important design and implementation implications for PD programs for the EFL teachers of Uzbekistan. Based on these problems identified, PD programs must tackle teacher provision with concrete strategies on various classroom management techniques and the use of technology, as well as increasing language proficiency. Furthermore, PD programs need to be crafted to address the specific needs of Uzbekistani EFL teachers in order to ensure that the professional development they receive will have immediate relevance to their classroom settings. Follow-up support, either mentoring or peer collaboration, is also crucial to facilitate the sustainability of professional development.

Lastly, this study contributes to a better understanding of EFL teacher professional development needs in Uzbekistan and provides valuable insights for developing more effective, context-based PD programs. By addressing the distinctive needs of teachers and

aligning PD programs with professional development needs, it is possible to enhance English language teaching quality in Uzbekistan alongside in-service teachers' continuous professional development.

References

- Abdullaev, R. (2021). *The Impact of National Independence and Globalization on the Status of English Language Instruction within Uzbekistan*. BCES Conference Books, 2021, Volume 19. Sofia: Bulgarian Comparative Education Society. ISSN 2534-8426, ISBN 978-619-7326-11-6
- Adey, P. (2004). The Professional Development of Teachers: Practice and Theory. In *Kluwer Academic Publishers eBooks*. <https://doi.org/10.1007/0-306-48518-4>
- Akhadaliyevich, A. K., & Nozimakhon, G. (2022). The role of needs analysis in language teaching. In the International Multidisciplinary Conference, "Innovative Research on the Modern Scientific World (pp. 67-71).
- Alkhateeb, H., Romanowski, M. H., Sellami, A., Abu-Tineh, A., & Chaaban, Y. (2022). Challenges facing teacher education in Qatar: Q methodology research. *Heliyon (London)*, 8(7), e09845. <https://doi.org/10.1016/j.heliyon.2022.e09845>.
- Asian Development Bank. (2022). *Enhancing teacher development in Central Asia: Key strategies for success*. Manila: ADB.
- Bezborodova, A., & Radjabzade, S. (2022). English in higher education in the Kyrgyz Republic, Tajikistan, and Uzbekistan. *World Englishes*, 41(1), 72-91.
- Cavendish, W., Barrenechea, I., Young, A. F., Díaz, E., & Avalos, M. (2021). Urban teachers' perspectives of strengths and needs: The promise of teacher-responsive professional development. *The Urban Review*, 53(2), 318-333.
- Dunn, K., & Iwaniec, J. (2023). English Impact: An Evaluation of English Language Capability: Uzbekistan. Education Sector Plan (ESP) of Uzbekistan 2019–2023, <https://www.globalpartnership.org/sites/default/files/2019-04-gpe-esp-uzbekistan.pdf> (accessed: 1.09.2020).
- English reform project | British Council*. (n.d.). <https://www.britishcouncil.uz/en/teach/reform-Zroject#:~:text=The%20new%20pre%2Dservice%20teacher,provide%20all%20future%20teachers%20>
- English Speaking Nation: Secondary Teacher Training (ESN: STT) — AC Uzbekistan*. (n.d.). AC Uzbekistan. <https://uzbekistan.americancouncils.org/esn-stt-en#:~:text=Implemented%20by%20the%20American%20Councils,in%2Dperson%20and%20in%2Dclassroom>
- Eshchanova, R., Bekchanova, D., & Bobojonova, G. (2020). The Current Core Of Education Reforms In Uzbekistan: One Step Forward Two Steps Back?. *European Journal of Research and Reflection in Educational Sciences*, 8(2056-5852), 169-182.
- Farrell, T. S., & Jacobs, G. M. (2020). *Essentials for successful English language teaching*. Bloomsbury Publishing.

Hutchinson, T., & Waters, A. (1987). English for specific purposes. Cambridge university press.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language teaching*, 39(2), 83-101.

Ibragimov, A., & Muradov, I. (2022). Based On Foreign Experiences: Teacher Training in Uzbekistan Improvement Issues. *Science and Innovation*, 1(7), 1128-1137.

Kosimov, A. A. (2021). The importance of needs analysis in teaching and enhancing English language proficiency among Uzbek EFL learners. *Asian Journal of Research in Social Sciences and Humanities*, 11(11), 616-623.

Kosimov, A., & Latipov, S. (2022). The importance of language education policy and practice in state education systems. *Involta Scientific Journal*, 1(9), 23-27.

LEX.UZ—Decree by the Cabinet of Ministers of Uzbekistan. (2022). № 01/19/34, 2022 <https://lex.uz/docs/-5831978>

LEX.UZ - Presidential Decree PF-79-сoн 26.05.2023. *Maktabgacha va maktab ta'limi vazirligi hamda uning tizimidagi tashkilotlar faoliyatini samarali tashkil etish chora-tadbirlari to'g'risida*. (n.d.). <https://lex.uz/docs/-6476175>

LEX.UZ - Presidential Decree PQ-5117-сoн 19.05.2021. *O'zbekiston Respublikasida xorijiy tillarni o'rganishni ommalashtirish faoliyatini sifat jihatidan yangi bosqichga olib chiqish chora-tadbirlari to'g'risida*. (n.d.). <https://lex.uz/docs/-5426736>

LEX.UZ - Presidential Decree PQ-4963-сoн 25.01.2021. *Xalq ta'limi sohasidagi ilmiy-tadqiqot faoliyatini qo'llab quvvatlash hamda uzluksiz kasbiy rivojlantirish tizimini joriy qilish chora-tadbirlari to'g'risida*. (n.d.). <https://lex.uz/docs/-5239538>

LEX.UZ - Decree by the Cabinet of Ministers of Uzbekistan. #25-сoн 17.01.2022. *Xalq ta'limi xodimlarini uzluksiz kasbiy rivojlantirish tizimini tashkil etish tartibi to'g'risidagi nizomni tasdiqlash haqida*. (n.d.). <https://lex.uz/ru/docs/-5828765?ONDATE=18.01.2022%2000>

LEX.UZ - Decree by the Cabinet of Ministers of Uzbekistan. (2021). № 05/19/312, 2021 <https://lex.uz/ru/docs/5431845>

Long, M. H. (Ed.). (2005). Second language needs analysis. Cambridge University Press.

Mamadaminova, N., & Madalińska-Michalak, J. (2025). Continuing professional development of teachers in Uzbekistan: From challenges to opportunities. *Pedagogical Quarterly/Kwartalnik Pedagogiczny*, 70(1).

Mamadaminova, N., & Khadjikhanova, S. (2023). Navigating continual progress: Insights into teacher professional development in Uzbekistan. *Pedagogical Quarterly/Kwartalnik Pedagogiczny*, 68(4).

Ministry of Education, Uzbekistan (2022) "10,000 Schools in Basic Education" project.

Mukhiddinova, E., & Agzamkhodjaeva, S. (2025). Effective Strategies for Teaching Large Classes: Challenges and Solutions. *Imras*, 8(5), 44-47.

Nazirova, S. O., Jalolova, S. M., Agzamova, D. B., Mamatova, F. M., & Yusupova, S. B. (2023). The problems faced by the teachers in teaching English as a foreign language in Uzbekistan. *Journal of Law and Sustainable Development*, 11(12), e2698-e2698.

Richards, J.C., Farrell, T.S.C., 2005. Professional Development for Language Teachers. Cambridge University Press, New York.

Richterich, R., & Chancerel, J. L. (1978). Identifying the Needs of Adults Learning a Foreign Language.

Sadeghi, K., & Richards, J. C. (2021). Professional development among English language teachers: challenges and recommendations for practice. *Heliyon (London)*, 7(9), e08053. <https://doi.org/10.1016/j.heliyon.2021.e08053>

Shaturaev, J. (2022). A case study: insights from the public education system of Uzbekistan. *Academic Research in Educational Sciences Volume 3 | Issue 8* ISSN: 2181-1385.

Shaturaev, J. (2021). A comparative analysis of public education systems of Indonesia and Uzbekistan. *Bioscience Biotechnology Research Communications*, 14(5), 89-92.

Shavkatova, Z., & Mirzaahmedov, M. (2025). Professional Development of Foreign Language Teachers: The Role Of Social Media In Professional Development. *Наука и инновация*, 3(16), 52-56.

Tashkent State University of Pedagogy. (2021). *Analysis of teacher training programs in Uzbekistan: Current state and future needs*. Tashkent: TSUP.

Turaeva, I. (2025). *Student's Academic Performance In Large-Sized Groups In Educational System Of Uzbekistan*. *Академические исследования в современной науке*, 4(8), 75-81.

Uzbekistan - Second Basic Education (English) (2015). Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/697331475117120787/Uzbekistan-Second-Basic-Education>.

West, R. (1994). Needs analysis in language teaching. *Language teaching*, 27(1), 1-19.

Appendix A - Participant Consent Form (Sample from one P.)



Informed Consent Form

Study title: Designing and Evaluating the Effectiveness of a Need-Based Professional Development Program (PDP) for In-Service EFL Teachers in Uzbekistan

Researcher: Munisa G. Muhtarova

Contact: munisismile1@gmail.com, m.muhtarova@tmuni.org, gs70000@upm.edu.my

1) Purpose of the study

You are invited to participate in a research study that explores (a) teachers' professional development needs, (b) the design and implementation of a need-based PDP, and (c) the PDP's impact on teaching practice.

2) What participation involves

If you agree, you may be asked to take part in the following activities:

- Complete a short demographics form
- Participate in a Focus Group Discussion (FGD) and/or interview (audio-recorded)
- Implement PDP activities and submit teaching artifacts (e.g., lesson plans/materials you create)
- For a small subgroup: one classroom lesson will be observed, followed by a short one-on-one interview (audio-recorded)

Estimated time: FGD/interview: 45 minutes; observation: one lesson; post-observation interview: 30 minutes.

3) Audio recording

FGDs and interviews will be audio-recorded only to ensure accurate transcription and analysis. Recordings will not be shared outside the research team.

4) Classroom observation



Observation focuses on teaching practices and classroom routines (not evaluating you as an employee). No video will be taken, but one or two snapshots of the lesson will be taken. Student names will not be recorded.

5) Voluntary participation and right to withdraw

Your participation is voluntary. You may refuse any question or activity and may withdraw at any time without penalty. If you withdraw, your data will be removed from the study where feasible (unless it has already been fully anonymized and integrated into the analysis).

6) Confidentiality

Your identity will be protected using codes (e.g., P1, P2). Reports will not include your name or school name. Any quotations used will be anonymized. Data will be stored securely (password-protected digital storage) and accessed only by the researcher/supervisory team.

7) Risks and benefits

Risks are minimal (possible discomfort when discussing challenges). You may benefit from reflecting on your practice and receiving PD support through the PDP.

8) Consent (please tick)

I confirm that I have read and understood the information above and agree to participate.

☒ I consent to participate in the study.

☒ I consent to audio recording of FGDs/interviews.

☒ I consent to classroom observation (one lesson) if selected.

☒ I consent to the use of my teacher-created materials (lesson plans/artifacts) for research purposes (anonymized).

Participant Code:

Signature:

[Handwritten Signature]

Date:

30.05.25

Researcher name: Munisa G. Muhtarova

Signature:

[Handwritten Signature]

Date:

30.05.25

Appendix B - Ethics Exemption Letter



116, Usta Shirin St., Tashkent, Uzbekistan
Tel: +998 712000123, info@tmuni.org

Tashkent Metropolitan University
School of Arts and Humanities

Reg #: 19/25

Date: 10th April, 2025

ETHICS EXEMPTION / WAIVER LETTER

To Whom It May Concern,

This letter confirms that the research project titled

“Designing and Evaluating the Effectiveness of a Need-Based Professional Development Program for In-Service English as a Foreign Language (EFL) Teachers in Uzbekistan”

Conducted by Munisa Muhtarova, Associate Professor, School of Arts and Humanities, TMU, it was reviewed at the institutional level for ethical considerations.

Based on the information provided, the project is classified as minimal-risk educational research involving adult participants (in-service teachers) and does not include medical interventions or vulnerable groups. Therefore, the project qualifies for an ethics review exemption/waiver under institutional practice, provided that the following conditions are met:

1. Participation is voluntary and based on written informed consent.
2. Participant confidentiality is protected through anonymization (no names, no identifying school information in publications).
3. Data (audio recordings, transcripts, and questionnaires) are stored securely and used solely for research purposes.
4. Participants may decline to answer any question or withdraw without penalty.

This exemption/waiver is granted for data collection activities conducted in the Olmazor District, Tashkent.

If the research design changes significantly (e.g., inclusion of minors, collection of sensitive personal data, medical procedures, or experimental interventions), additional ethics review will be required.

Sincerely,

Bekhzod Djalilov,
TMU President

Signature: _____



Appendix C - Open-ended questionnaire Protocol

Stage 1: Collecting a demographic data

OPEN-ENDED QUESTIONNAIRE

Please answer the following questions to the best of your knowledge and ability. Responses will be evaluated for correctness, comprehensiveness, and clarity.

1. What's your nationality

2. Marital Status

- Married
- not married

3. Teaching experience:

- 1-5 years
- 6-10 years
- 11 years and above

4. Age

- 21-25
- 26-30
- 31-35
- 36-40
- Over 40

5. What is the highest level of education and field of study?

- Bachelor's Degree
- Master's Degree

6. Majoring in a program?

7. Year of graduation?

8. What grades do you teach?

9. Have you taken any of the following language proficiency tests? What was your score?

- ELTS _____

- CEFR _____
- Duolingo _____
- APTIS for teachers _____
- Other (please specify) _____
- Never taken a test _____

10. Have you attended any teacher training workshops, seminars, or webinars related to teaching English before? Please specify.

11. Your position at school (example: senior teacher) _____

Appendix D - FGD Interview Protocol

Instrument #2- 14 questions

5 FGD groups with 5 members in each.

Purpose: to identify in-service EFL teachers' priority professional development needs, barriers to transferring PD into classroom practice, and preferred features of an effective, need-based PDP, informing RQ1 and guiding RQ2 design decisions.

Note: Probes were used only when needed to encourage concrete examples from the participants.

1. What was the most recent PD program you attended like?
Probes: lecture vs practice, relevance, what felt useful.
2. After PD ends, what happens—do you apply it in class? Why or why not?
Probes: workload, time, leadership expectations, confidence, lack of follow-up.
3. Describe a time you tried to apply PD learning and it didn't work. What blocked the transfer?
Probes: class size, discipline, mixed levels, materials, unclear steps, no coaching.
4. What PD format helps you transfer learning into practice quickly?
Probes: demo lessons, rehearsal/micro-teaching, ready-to-use routines, peer support.
5. What are your top three daily challenges in your English classes?
Probes: mixed proficiency, motivation, discipline, exam pressure, and L1 reliance.
6. What limits your lesson planning and teaching quality in a typical week?
Probes: paperwork, teaching load, admin tasks, resources, stress.
7. How mixed are student levels in one class, and how do you currently handle it?
Probes: who gets left behind, current strategies, and what support is missing.
8. If you could choose only three PDP topics that would make the greatest difference, what would they be and why?
Probes: management routines, differentiation, speaking, assessment/feedback, listening, and writing.
9. Which teaching tasks do you feel least confident about right now?
Probes: group work control, speaking facilitation, monitoring, feedback, and pacing.

10. In your context, what works better—communicative activities or traditional grammar-focused teaching? Why?

Probes: student response, control, time, exams, and teacher language confidence.

11. What worries you most when you try pair/group speaking activities?

Probes: noise, losing control, students using Uzbek, weak students, time.

12. What technology do you realistically have access to at school, and what barriers exist?

Probes: internet, projector/TV, printing, phone policy, electricity.

13. If PD used an LMS/online platform, what would make it feasible for you?

Probes: mobile-friendly, low data use, short tutorials, printable alternatives, optional use.

14. What must an “ideal PDP” include so you keep using it after the program ends?

Probes: follow-up, coaching, peer triads, evidence/reflection tools, and ready-to-print packs.

Appendix E - Codebook excerpt (FGDs thematic analysis)

Dataset: 5 FGDs (25 participants), 5 June 2025 (English)

Software: Quirkos (coding/retrieval) + Microsoft Excel (codebook/audit trail)

Coder: Researcher (single-coder)

Theme mapping:

- Theme 1 (Large/mixed-ability classes): CM1–CM3
- Theme 2 (Technology constraints): TECH1–TECH3
- Theme 3 (Speaking confidence & communicative teaching): LANG1–LANG2, PED1
- Theme 4 (Practical/context-specific PD): PD1–PD3
- Theme 5 (Continuous support): SUP1–SUP3

Code ID	Code label	Definition	Include when...	Exclude when...	Exemplar quotation
CM1	Large class size constraint	Limits created by high student numbers for participation/monitoring	Teacher links difficulty to class size	Mentions mixed levels but no size reference	“It is difficult to make all of them active if you have 35 students in a class.” — P3, FGD Group 1
CM2	Mixed-ability challenge	Difficulty teaching students with different proficiency levels in one class	Teacher contrasts advanced vs low-level learners	Only discipline/behavior without level differences	“I have students who are advanced, and others who are elementary level.” — P3, FGD Group 1
CM3	Need for differentiation strategies	Expressed need for practical ways to handle differences simultaneously	Teacher states inability to manage differences effectively	Pure complaint without need/strategy gap	“I do not know how to manage differences effectively.” — P3, FGD Group 1
TECH 1	Lack of equipment/infrastructure	Absence of devices/ICT facilities needed for tech use	Mentions no equipment, computers, internet, etc.	Skill/confidence only without resources	“We just do not have the equipment.” — P7, FGD Group 2
TECH 2	Low tech-integration capacity	Limited know-how to use tech pedagogically even if available	Says they wouldn’t know how to use tools while teaching	Only resource constraints	“Even if we had computers, we would need some training on teaching tools/applications.” — P7, FGD Group 2

TECH 3	Desire for web/digital resources	Interest in using online resources/tools	Expresses a wish to use web resources	Complaints only, with no expressed desire	"I would like to use more web resources..." — P7, FGD Group 2
LAN G1	Speaking anxiety / low confidence	Nervousness about speaking English in front of learners	Mentions nervousness/ anxiety about speaking	Grammar confidence only	"If I have to speak English fluently in front of students, I feel nervous." — P10, FGD Group 3
LAN G2	Grammar–speaking imbalance	Perceived strength in grammar but weakness in speaking	Contrasts grammar confidence with speaking difficulty	Speaking difficulty without grammar comparison	"I am confident in grammar, but..." — P10, FGD Group 3
PED1	Difficulty teaching speaking	Difficulty facilitating communicative speaking instruction due to teacher's confidence	Links teacher confidence to teaching speaking	Student speaking difficulty only	"It's hard to teach speaking if you're not confident yourself." — P10, FGD Group 3
PD1	PD not practical / too theoretical	PD was described as not providing actionable classroom steps	Says workshops don't teach "how to" in practice	Critiques unrelated to practicality	"They don't really instruct me on how to manage my classroom or integrate technology in practice." — P12, FGD Group 4
PD2	Need for realistic, context-specific PD	PD should match teachers' real constraints and context	Mentions "realistic," "context-specific," "our situation."	Generic desire for PD without context	"I need training that is more realistic and context-specific in our situation." — P12, FGD Group 4
PD3	PD should address core classroom problems	PD should target classroom management and tech integration as priority needs	Mentions classroom management and tech integration as PD needs	Needs that are purely language proficiency only	"Manage my classroom or integrate technology in practice." — P12, FGD Group 4
SUP1	No follow-up after PD	PD ends with no continuing support	Mentions no follow-up after workshops	Mentions follow-up exists	"When they're over, there's no follow-up." — P15, FGD Group 5

SUP2	Need for mentoring/coaching	Desire for accountability/support (mentor, checking, guidance)	Mentions someone checking/reminding/supporting implementation	General “support” without a mechanism	“You can’t do everything without someone checking on you or reminding you where you stand.” — P15, FGD Group 5
SUP3	Transfer support (implementation help)	Need for help applying PD into practice	Links follow-up to classroom transfer	Complaints about PD content only	“We attend these workshops, yet... there’s no follow-up.” — P15, FGD Group 5

Appendix F - Member check

Member Checking Validation

Instructions:

Please read each theme summary. For each theme, tick the option that best reflects your view and add comments if needed. You may write in English or Uzbek.

Participant code: FGD 3; P 4

Date: June 10

Theme 1: Large Mixed-Ability Classes

Does this reflect your experience? ☒ Yes ☐ Mostly ☐ Partly ☐ No

Comments: It is very hard to teach and control large classes.

Theme 2: Technology in Teaching

Does this reflect your experience? ☒ Yes ☐ Mostly ☐ Partly ☐ No

Comments: We often have limited or broken technology and weak internet, it is hard to use videos or online tools in our lessons.

Theme 3: Speaking confidence & communicative practice

Does this reflect your experience? ☒ Yes ☐ Mostly ☐ Partly ☐ No

Comments: Many students are shy to make mistakes, so they need more speaking practice to build confidence.

Theme 4: Practical, context-specific PD

Does this reflect your experience? ☒ Yes ☐ Mostly ☐ Partly ☐ No

Comments: We need trainings that are practical, and match our classroom situation. We need simple activities which we can use right away.

Theme 5: Lack of Support for Continuous Professional Development

Does this reflect your experience? ☒ Yes ☐ Mostly ☐ Partly ☐ No

Comments: After trainings there is little time, guidance or support to continue learning. It is hard to improve step by step.

Overall question:

Are there any significant needs, challenges, or experiences that are missing from these themes?

☒ No ☐ Yes (please explain): Everything seems to be covered.

Appendix G - Prototype Needs-Based PDP Model

Facilitator/Trainer: The researcher—being an expert in EFL pedagogy and teacher development—led each and every session of professional development by offering the program in a practicing format adjusted to local classroom realities.

Module (linked to needs/themes)	Learning outcomes (participants will be able to...)	Delivery mode & core activities	Monitoring & evaluation indicators (examples)
1. Managing large, mixed-ability classes	Apply low-prep differentiation routines (tiered tasks, flexible grouping, scaffolding) and implement practical classroom-management routines that support participation and discipline in large classes.	Demonstration lesson—guided deconstruction - micro-teaching rehearsal - peer feedback using checklist - classroom-ready lesson plan.	Pre/post self-efficacy (classroom management); observation checklist (use of routines); teacher reflection log (what was applied).
2. Low/no-tech teaching (“minimum viable tech”)	Adapt textbook tasks into interactive no-tech activities and use realistic low-tech options (e.g., phone audio/offline tools) where feasible.	Task-adaptation workshop—creation of activity bank—try-out and troubleshooting—sharing of printable templates.	Review of produced artifacts (activity bank/templates); implementation log; short follow-up check-in survey on feasibility.
3. Teaching speaking + teacher speaking confidence	Use classroom-English routines confidently and facilitate structured speaking activities (pair work, information gap, role-play) with clear steps and monitoring.	Micro-teaching cycles—coached feedback - classroom-English scripting - rehearsal of speaking-task facilitation.	Pre/post confidence rating; micro-teaching rubric; observation checklist for speaking-task implementation.
4. Communicative methods under constraints	Convert grammar-heavy lessons into communicative practice sequences and apply strategies to engage low-motivation learners in mixed-level classrooms.	Model lesson—lesson redesign sprint—peer review—implementation planning (“next-week plan”).	Lesson plan quality rubric; teacher implementation log; student engagement proxy (participation tally).

5. Transfer support: follow-up and community of practice	Complete one peer observation and feedback cycle and participate in a PLC circle to troubleshoot implementation barriers.	PLC circles—peer observation protocol—reflection discussion—problem solving clinic.	PLC attendance; completed observation forms; reflection notes; follow-up survey on transfer barriers/support needs.

Between Face-To-Face, Hybrid and Online: How Shifting Modalities Shape Rapport and Communicative Competence Among Malaysian University Students

ABSTRACT

Authors: Siti Nadhirah Abd Rahman^{1*} & Abu Bakar Razali¹

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
*Corresponding author:
s_nadhirah@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Developing communicative competence in second language learners is challenging due to limited exposure and opportunities to use English, challenges that were heightened by the COVID-19 pandemic. The abrupt shift to online and hybrid learning disrupted students' ability to build rapport and engage in meaningful interactions, with reduced non-verbal cues, spoken communication, personalised feedback, and outside-classroom practice contributing to misunderstandings, isolation, and lower participation. Language learning is closely tied to social context, which fosters socialisation and the development of communicative competence. This exploratory qualitative study examines how classroom modality, face-to-face, hybrid, and online, shapes rapport development among ESL university students and how these experiences shape learners' perceived opportunities to engage in classroom communication that supports the development of communicative competence. In-depth semi-structured interviews with 20 university students revealed that there was a shift in individual interactions, class engagement and participation, asking and addressing questions, the role of spoken communication, after-class bonding, and non-verbal communication in online and hybrid learning environments. The findings offer pedagogical insights for ESL instructors and inform higher education stakeholders seeking to design interactionally supportive learning environments across instructional modalities.

Keywords: Communicative Competence, ESL Learners, Higher Education, Second Language Acquisition.

1. Introduction

Communicative competence enables language learners to communicate clearly, effectively, and appropriately, helping them avoid misunderstandings in everyday interaction (Hymes, 1972). However, for second language learners, developing this competence is far from straightforward. Many experience limited exposure to the target language and few opportunities to use it authentically, and these challenges became even more pronounced during the Covid-19 pandemic, when education systems around the world were abruptly forced to transition online.

1.1. Research Background and Problem

With the rapid spread of Covid-19, physical interaction came to a halt. A systematic review by Liu, Wyver, and Chutiyami (2022) highlights how pandemic restrictions disrupted children's outdoor activities and socialisation. Research has shown that such disruptions affected not only communication quality but also learners' opportunities for meaningful interaction and socialisation (Long et al., 2022; Towner et al., 2022). In language learning contexts, where interaction plays a central role in developing communicative competence, these changes were particularly consequential. Students often experienced reduced participation and engagement in virtual classrooms, partly due to the absence of physical co-presence and the ease with which learners could disengage from online interaction (Singh, Steele, & Singh, 2021). These circumstances highlight the importance of examining how shifts in instructional modality influence interpersonal dynamics and communication in ESL classrooms.

Although online learning is not a new concept given the long history of distance education, its success depends on factors such as maintaining meaningful teacher–student interaction, supporting authentic learning tasks, ensuring engagement, and designing effective communication structures (Agostinelli, 2019). During emergency remote teaching, many of these factors were difficult to achieve. Students could easily disengage by switching off their cameras and this affected their willingness to participate and interact (Singh, Steele, & Singh, 2021). Much of the existing literature focuses on the challenges faced by educators in adapting to online teaching, leaving a gap in our understanding of students' lived learning experiences during this period (Peimani & Kamalipour, 2021). This gap is particularly concerning in ESL contexts, where social interaction forms the foundation of language learning.

Halliday (1978, 2004) emphasises that language development occurs through socialisation, where learners acquire both language and meaning from interactions within their environment. When physical socialisation became restricted, this natural learning pathway was disrupted. Studies have reported decreased communication, reduced motivation, and heightened feelings of isolation among students in online ESL environments (Alawamleh et al., 2022; Long et al., 2021). These findings raise an important question: how have changes in modality, moving from face-to-face to online, then hybrid, and eventually back to face-to-face, affected ESL learners' ability to build rapport and develop communicative competence?

In navigating these challenges, ESL instructors adopted various strategies to sustain interaction in virtual spaces, such as synchronous computer-mediated communication (Junn, 2023), online intercultural exchanges (Taskiran, 2023), and task-based pedagogical adaptations (Belda-Medina, 2021). While these approaches aimed to emulate authentic communication and support

learners' motivation, the long-term effects of shifting modalities on students' socialisation habits remain unclear.

Rapport, a mutual, prosocial, and trusting connection between individuals (Frisby & Martin, 2010), is particularly relevant here. Rapport plays a crucial role in language learning because it shapes comfort, participation, and learners' willingness to communicate. Research shows that different teaching modalities can significantly influence rapport-building experiences (Denker et al., 2022), which in turn affect participation and learning outcomes (Frisby & Martin, 2010).

In light of this, the present study aims to explore how modality shifts, specifically face-to-face, hybrid, and online learning, shape rapport development in the ESL classroom and how these shifts may influence learners' communicative competence. The university students chosen for this study had enrolled in a compulsory Academic Interaction course taken in their first year that explicitly focused on the development of students' communicative competence. By foregrounding students' experiences of rebuilding rapport after prolonged online learning, the study provides pedagogical insights for ESL instructors and practical considerations for higher education policy and curriculum planning units in designing learning environments that support meaningful interaction and communicative development across modalities.

For the purposes of this study, rapport refers to the relational quality that fosters comfort, trust, and interaction enjoyment between learners and instructors within the classroom environment (Frisby & Martin, 2010). Such rapport is reflected in learners' willingness to participate, share ideas, and engage in communicative activities with the instructor and other learners.. Communicative competence is conceptualised following Hymes' (1972) framework, encompassing grammatical, sociolinguistic, and strategic competence. These constructs informed the design of the semi-structured interview prompts used in the study (see Appendix A), which explore students' experiences of classroom interaction, participation, and rapport-building across different instructional modalities.

1.2. Research Purpose

This study adopts an exploratory qualitative approach, as empirical research examining rapport development across shifting instructional modalities in ESL contexts remains limited. It is guided by a single overarching research aim, within which the analysis focuses on learners' experiences of rapport development across face-to-face, hybrid, and online classrooms, their perceptions of rapport in supporting communicative competence, and factors influencing the rebuilding of rapport after prolonged online learning.

Guided by this overarching purpose, the analysis attends to the following interrelated aspects of learners' experiences:

1. Rapport development across face-to-face, hybrid, and online ESL classroom settings
2. Learners' perceptions of the role of rapport in supporting communicative competence following modality shifts
3. Factors that facilitate or hinder the rebuilding of rapport after prolonged online learning

2. Literature Review

2.1. Impact of Online Modalities on Socialisation and Rapport

Non-verbal cues such as facial expressions, gestures, and body language are central to developing rapport and interpreting meaning. In online and hybrid settings, these cues are often reduced or lost entirely, which can lead to misunderstandings, especially among ESL learners who rely heavily on contextual clues (Ying et al., 2021; Nesprava et al., 2023). Although online platforms offer alternative modes of interaction, they rarely replicate the richness of in-person communication (Ramalingam et al., 2023; Hamid et al., 2025). Students frequently report reduced communication and feelings of isolation when learning online (Alawamleh et al., 2020), and such experiences inevitably affect their rapport with instructors and peers (Singh et al., 2021). Existing studies primarily document the technological limitations of online communication, while offering limited insight into how learners themselves experience these changes in rapport-building across different instructional modalities.

Another challenge lies in the reduced personalised interaction between lecturers and students. Many learners describe feeling disconnected from instructors in online environments, which affects the relational quality essential for rapport-building (Rahman & Razali, 2024; Ash'ari et al., 2025). Even when educators utilise multimodal tools or adapt their teaching approaches, these interactions remain mediated through screens, and the immediacy normally present in face-to-face classrooms is diminished (Belda-Medina, 2021). For ESL learners who depend on scaffolded support to develop communicative competence, this relational distance may constrain both engagement and language use opportunities.

Asking questions is a natural part of language learning, yet many students find it less comfortable in online environments. The absence of non-verbal encouragement and delayed instructor responses often discourage learners from seeking clarification (Hamid et al., 2025; Nesprava et al., 2023). Some hesitate out of fear of miscommunication, while others remain silent due to uncertainties about online etiquette or response time. This hesitation may further reduce opportunities for immediate feedback and interactive negotiation of meaning, both of which are essential processes in second language learning. These changes have been linked to reduced cognitive engagement (Peimani & Kamalipour, 2021; Towner et al., 2022), signalling broader implications for ESL participation.

2.2. Shifts in Classroom Communication Dynamics

Spoken interaction forms a central component of ESL learning, as it allows learners to practise language use and negotiate meaning in real time. Studies indicate that spoken communication tends to be less prominent in online and hybrid classrooms, where instructional time is often structured around content delivery rather than spontaneous discussion (Singh et al., 2021; Rahman & Razali, 2024). Although technological tools provide structure, they still cannot mirror the organic flow of real-time conversation (Ramalingam et al., 2023). Studies show that online environments often prioritise content delivery rather than verbal interaction, which may limit opportunities for learners to practise speaking (Singh et al., 2021; Rahman & Razali, 2024). Current research focuses largely on instructional practices rather than examining how learners perceive these changes in interactional opportunities and how such experiences influence rapport and communicative engagement.

Full-class discussions also take on a different character online. Instead of dynamic, free-flowing exchanges, discussion tends to become segmented and less inclusive (Ramalingam et al., 2023). These limitations can hinder the development of collective communicative competence (Roinah et al., 2024). Alawamleh et al. (2020) also note that online classes restrict robust discussions, while Long et al. (2021) highlight how the reduction in social interaction erodes relational capital within the classroom community.

2.3. The Decline in Outside-Classroom Interactions

Outside-classroom interaction plays a significant role in strengthening communicative competence, particularly for ESL learners who benefit from informal language practice. Pandemic restrictions significantly reduced these opportunities, leading to fewer peer interactions and less authentic language use (Hernandez-Ledezma & Janecek, 2025). Towner et al. (2022) further explain that reduced socialisation contributed to increased loneliness, which affects learners' confidence and motivation to communicate. When informal learning opportunities diminish, the development of communicative competence becomes even more challenging. Despite these observations, relatively little research has explored how learners experience the loss and subsequent rebuilding of informal interaction as instructional modalities shift back toward hybrid or face-to-face learning environments.

2.4. Implications for Pedagogical Practice

The reduction of non-verbal cues in online and hybrid learning underscores the need for intentional strategies that promote meaningful oral interaction. Existing studies highlight techniques such as breakout rooms, structured oral tasks, and small-group discussions can recreate aspects of face-to-face engagement, addressing relational and communicative gaps in digital environments (Cho & Cho, 2016; Fehervari et al., 2022; Almendingen et al., 2022; Saltz & Heckman, 2020). However, much of this work remains instructor-centred or strategy-focused, with limited attention to learners' perceptions of rapport across instructional modalities. By exploring how modality shifts affect rapport and communicative competence, this study can inform which strategies are most effective in sustaining interaction and participation among ESL learners.

Scaffolding is equally crucial in online contexts, where limited individual support can lead to disengagement. Multimodal tools, such as interactive videos with embedded quizzes or collaborative online platforms, can enhance engagement while supporting learners' communicative development (Smallwood & Brunner, 2017; Bakla & Demiröz, 2024). While scaffolding through multimodal tools has been shown to support engagement (Smallwood & Brunner, 2017; Bakla & Demiröz, 2024), there is insufficient qualitative evidence examining how such practices influence rapport rebuilding and communicative competence from the learner perspective.

Finally, online learning can disrupt social cohesion, highlighting the importance of opportunities for peer collaboration and informal interaction (Ong & Quek, 2023). Peer-based tasks and interactive group activities can foster rapport and support language fluency, particularly for learners with fewer natural exposure opportunities (Smallwood & Brunner, 2017). By examining ESL students' experiences across modalities, this study can guide pedagogical practices

that strengthen social bonds and communicative competence, even in non-physical learning environments.

Notably, few studies have explicitly examined rapport as a key construct linking modality shifts to communicative competence in ESL higher education contexts. Addressing this gap, the present study explores ESL learners' experiences across face-to-face, hybrid, and online modalities, clarifying how rapport-related interactional behaviours shape communicative development following prolonged online learning.

2.5. Theoretical Framing

This study is conceptually informed by social-interaction perspectives on language development and by research on learners' willingness to communicate (WTC). From a Vygotskian perspective, language learning occurs through social interaction, where meaning is co-constructed through communicative activity and supported by interpersonal relationships within the learning environment (Vygotsky, 1978). In classroom contexts, opportunities for interaction with teachers and peers create the conditions through which learners practise language use and gradually develop communicative competence. Changes in instructional modality may therefore influence not only the frequency of interaction but also the relational dynamics that support communicative engagement.

Complementing this perspective, the willingness-to-communicate framework highlights how learners' readiness to initiate communication is shaped by situational and relational factors, including perceived comfort, confidence, and interpersonal rapport within the classroom (MacIntyre et al., 1998). In language learning contexts, rapport between instructors and students can foster a supportive environment that encourages learners to participate more actively in communicative tasks.

These theoretical perspectives informed the analytical approach adopted in this study. During the thematic analysis (Braun & Clarke, 2006), initial codes were developed around learners' interactional experiences, perceptions of relational comfort, and opportunities for participation across different instructional modalities. Themes were subsequently refined to capture how modality shifts influenced rapport development and the interactional conditions that shape learners' communicative engagement in ESL classrooms.

3. Methodology

3.1. Research Design

This study adopts an exploratory qualitative research design to examine ESL university students' experiences of rapport development and classroom interaction across different instructional modalities following the COVID-19 pandemic. An exploratory qualitative approach is appropriate when the aim is to gain an in-depth understanding of participants' perspectives and experiences in contexts where existing empirical evidence remains limited.

Data were collected through in-depth semi-structured interviews with undergraduate ESL students, allowing participants to describe their experiences of learning across face-to-face, hy-

brid, and online classroom environments. This approach enabled the researcher to capture nuanced insights into how students perceived rapport, interactional dynamics, and communicative engagement across these modalities.

The interview data were analysed using reflexive thematic analysis following Braun and Clarke's (2006, 2019) framework. This analytical approach allows patterns of meaning to be identified across participants' accounts while remaining sensitive to the contextual and experiential dimensions of the data. Through this process, the study aims to generate a rich understanding of how learners interpret their interactional experiences across different learning modalities and how these experiences shape the relational conditions that support classroom communication.

3.2. Participants and Setting

Participants were undergraduate ESL students enrolled in a university English language course that had experienced instructional shifts across face-to-face, hybrid, and online modalities. Participants were invited through an announcement disseminated by course lecturers following institutional ethical approval. Interested students registered their willingness to participate via a Google Form, which included screening questions to verify the study's inclusion criteria.

A purposive recruitment strategy was employed to ensure participants had relevant experience with the phenomenon under investigation. Specifically, participants were required to have experienced face-to-face instruction and at least one online or hybrid learning phase within the same course. Thirty students initially indicated interest through the form and met the inclusion criteria. However, ten students did not respond to follow-up communication for interview scheduling. The final sample therefore consisted of 20 participants who met the inclusion criteria and completed the interviews.

The inclusion criteria required participants to:

- Be enrolled in the selected ESL course, Academic Interaction and Presentations during the period of modality shifts,
- Have experienced face-to-face and at least one online or hybrid learning phase, and
- consent to participate in a recorded interview.

The study's participants were 20 university undergraduate students enrolled at a Malaysian public university. Participants were selected based on their enrolment in a compulsory course, the Academic Interaction course taken in their first year that explicitly focused on the development of students' communicative competence. The course entailed them to learn and be tested on how to communicate in group interactions as well as during presentations for 3 hours a week, throughout a duration of 14 weeks.

Each participant was identified with the pseudonym R1 – R20. The participants were informed that demographic data on them that did not impact the findings of the study would not be included in any publications or reports. All participants had experienced face-to-face, hybrid, and online modes of learning English as a second language.

After each interview session, the data was transcribed and coding was done. These codes were categorised and matched to develop themes. Data collection continued until data saturation was achieved, which was after the 20 university students were interviewed.

To assess the adequacy of the dataset, saturation was monitored during the coding process. As interviews were analysed sequentially, the emergence of new codes was tracked in relation to the study's research questions. After coding the later interviews (approximately interviews 18–20), no additional codes relevant to the research aims were identified, indicating that thematic saturation had been reached. This approach aligns with qualitative research guidance suggesting that saturation may be evidenced when additional data no longer contribute substantially new conceptual insights (Guest, et al., 2006).

3.3. Data Collection and Trustworthiness

The primary data collection method consisted of in-depth, semi-structured interviews. This method was chosen to allow participants to share detailed, rich narratives of their experiences, moving beyond simple responses (Merriam & Tisdell, 2016).

The interview protocol was developed and adapted based on the themes presented in Frisby and Martin's (2010) Modified Rapport Measure. While the original measure focused on quantifying rapport, the core constructs of comfort, relational quality, and interaction enjoyment were utilised to frame the open-ended interview questions. The initial 11 quantitative items were transformed into nine open-ended, non-leading questions to explore the following areas:

- Perceived comfort in communication with the instructor and peers
- Perceived changes in the student-instructor and student-peer relationships across different modalities
- Perceived impact of these relationships on the development of communicative competence

The finalised questions in the interview protocol can be found in Appendix A. The interviews lasted approximately 30 to 60 minutes, depending on the participants' responses. All interview questions were asked in English, and most of the responses were also in English, except when the participants were at loss of words and needed to revert to their mother tongue. In situations where the participants did not use English, the interviewer would attempt to rephrase the response in English to the participants, asking them if this is what they had meant earlier.

The interviews were conducted in person by a trained research assistant who was not involved in the participants' course assessment to minimise potential power dynamics. The research assistant was a postgraduate student with prior experience in qualitative interviewing and received training from the researcher regarding the study objectives, interview protocol, and appropriate probing techniques. This training included familiarisation with the research questions, discussion of relevant literature, and guidance on maintaining consistency while allowing participants to elaborate on their experiences.

All interviews were audio-recorded with participants' consent and transcribed verbatim immediately following completion. The interviews were conducted in-person by the research assistant in the campus area. The research assistant was trained, specifically through explanation of

the objectives of the study, the previous literature involved, as well as suggested manners for probing and redirecting the participants if needed. Prior to data collection, ethical approval was obtained from the Ethic Committee for Research Involving Human Subject (JKEUPM) under Universiti Putra Malaysia, reference number JKEUPM-2024-058.

The transcribed interview data were analysed using Thematic Analysis as outlined by Braun and Clarke (2006). The analysis was conducted using Atlas.ti to manage and organise the large dataset. To enhance the trustworthiness of the findings, member checking was conducted with participants, and peer debriefing was performed between the research assistant and researcher. The member checking was conducted within a duration of one to two weeks from the interview date. The participants were given a few days to review the transcripts and codes that were identified. None of the participants requested any changes to the data and coding. Once all the interviews and member checking procedures were completed, the researcher and research assistant conducted peer debriefing to discuss the findings and interpretations made, as well as the themes that could be derived from the data.

3.4. Data Analysis

Data were analysed using reflexive thematic analysis following Braun and Clarke's (2006, 2019) six-phase framework. Analysis began with repeated reading of interview transcripts to achieve familiarisation with the data.

The coding process followed several sequential steps. First, initial open coding was conducted inductively, focusing on participants' descriptions of rapport development, interactional behaviours, and communicative experiences across face-to-face, hybrid, and online modalities. These initial codes were then organised into a preliminary codebook, where each code was assigned a working definition based on patterns observed in the data.

Second, independent coding was conducted by two coders. The first coder (the research assistant) completed initial open coding across all transcripts, while the second coder (the researcher) independently coded the dataset using the same preliminary coding framework. This step allowed multiple perspectives to inform the interpretation of the data.

Third, the coders engaged in iterative consensus meetings to compare coding decisions, discuss interpretive differences, and refine the coding framework. During these discussions, several codes were merged, renamed, or clarified to enhance conceptual coherence. These collaborative discussions served to strengthen analytic transparency and ensure that the coding framework remained grounded in participants' accounts.

For example, one segment describing students' hesitation to speak during online classes was initially coded by the first coder as **"online participation anxiety,"** while the second coder interpreted it as **"lack of interactional cues."** During the consensus discussion, both coders reviewed the excerpt in context and agreed that the hesitation was primarily linked to the absence of non-verbal feedback from peers and instructors. The code was therefore revised to **"reduced non-verbal encouragement,"** which more accurately captured the interactional dynamic described by participants.

Finally, codes were organised into broader themes, and relationships among these themes were examined through the development of a thematic map, which guided the final interpretation of

patterns across the dataset. Code definitions were refined iteratively during this stage. An example of the final codebook, including one theme, its definition, and representative codes and excerpts, is provided in Appendix B.

Consistent with the principles of reflexive thematic analysis, formal inter-coder reliability statistics were not calculated, as coding agreement is not treated as a measure of analytic validity within this approach (Braun & Clarke, 2019). Instead, analytic rigour was maintained through reflexive dialogue between coders, peer debriefing, and grounding interpretations closely in participants' accounts.

Reflexivity was integrated throughout the research process to acknowledge potential researcher influence. The researcher is an ESL educator with experience teaching in face-to-face, hybrid, and online higher education contexts. While this positionality provided contextual sensitivity to rapport and interactional dynamics, it also required careful attention to avoid assuming shared experiences with participants. To address this, interpretations were continually revisited during coding discussions and member-checking procedures were conducted to ensure that findings reflected participants' perspectives.

All interviews were audio-recorded and transcribed verbatim by the research research assistant shortly after each interview session. The researcher subsequently reviewed the transcripts while listening to the recordings to verify accuracy and ensure that participants' responses were faithfully represented. Where minor transcription ambiguities occurred, they were resolved through discussion between the research assistant and the researcher. As the interviews were conducted in English, no translation was required.

To enhance the credibility of the findings, a member-checking procedure was conducted. Participants were contacted via email and provided with a brief summary of the preliminary themes derived from the analysis. They were invited to review whether the interpretations accurately reflected their experiences. Eighteen of the twenty participants responded to the member-check request, and none suggested changes to the interpretations presented. This process helped ensure that the analysis remained grounded in participants' perspectives.

To assess the adequacy of the dataset, saturation was monitored throughout the data collection and analysis process. Interviews were transcribed and preliminarily coded shortly after each interview, allowing the researcher to track the emergence of new codes in relation to the research questions. The coding framework began to stabilise after approximately the eighteenth interview, with subsequent interviews contributing additional examples and elaborations of existing codes rather than introducing substantively new categories. The remaining interviews were conducted to confirm the consistency of the themes across participants' experiences. This approach aligns with qualitative research guidance suggesting that saturation may be evidenced when additional data no longer generate substantially new conceptual insights (Guest, Bunce, & Johnson, 2006).

4. Findings and Discussion

This study aims to explore how the different modalities, namely face-to-face, hybrid, online, impact rapport in ESL classrooms for the development of students' communicative competence. The emerging themes from this study include the shift in individual interactions, class

engagement and participation, asking and addressing questions, the role of spoken communication, after-class bonding, and non-verbal communication.

4.1. Individual Interactions

Due to the changes in modality, the students in ESL classrooms experience a shift in individual interactions, with a feeling of distance and isolation due to the classroom related interaction being mostly formal and lacking the opportunity for building rapport.

Respondent 3 noticed a shift in interaction, “I do notice small changes depending on the mode of learning. In face-to-face classes, it's easier to connect personally, while in online or hybrid classes, communication can feel more formal”. Though formal language use is deemed appropriate for the classroom, there should also be opportunities for learners to experience individual and casual interaction, as it can be a medium for building rapport between the instructor and learners. Comparing to their experience in hybrid and online classrooms, Respondent 12 notes that “In face-to face class, I can feel a stronger more impactful relationship as there’s no barrier in communication”. Most of the students reported that they would feel that they experienced more personalised and individualised interactions in face-to face classes.

In online and hybrid learning environments, the personal touch often afforded by face-to-face engagement is challenging to replicate. Students frequently report feelings of isolation and disconnection from instructors, which can inhibit their ability to form meaningful relationships that are crucial for language acquisition (Rahman & Razali, 2024). Studies highlight the importance of personalized feedback and interactions in cultivating a conducive learning atmosphere (Ash’ari et al., 2025). This gap in individual attention may further exacerbate existing disparities in learner performance and confidence, particularly for those requiring more tailored support (Alaskar, 2023).

Respondent 18 also mentioned “I think the relationship with instructor will change especially through face-to-face because when we communicate in real life, we can feel the connection but if hybrid or online, it seems like there are some barriers to communicate because even we see through screen of the laptop but still we cannot feel their presence,” where they emphasize the lack of presence of the ESL instructor in hybrid and online mode, affecting communication in the classroom. This presence does not only refer to the physical presence, but also the salience of the ESL instructor, specifically in the personalised interactions. These personalised interactions are vital in students’ development of their communicative competence.

Alawamleh et al. (2020) discusses research where students felt less connection to their instructors in virtual settings, which hindered rapport-building crucial for language development. At the same time, Respondent 11 stated that “in online or hybrid settings, it can feel more distant but still effective if there’s regular interaction,” which does highlight the possibility for the communication in hybrid classrooms to be more effective if initiated constantly and consistently. Belda-Medina (2021) emphasizes this by advocating for more multimodal interactions to enhance communicative competence in an online framework, supporting the need for tailored approaches even in non-physical environments.

The findings show that changes in learning modality significantly affect individual interactions and, consequently, rapport in ESL classrooms. Students consistently report that face-to-face classes allow stronger, more personal connections due to the immediacy of communication and

the instructor's clearer presence, which supports the development of communicative competence. In contrast, hybrid and online environments tend to feel more formal, distant, and less personalised, often leading to feelings of isolation and reduced opportunities for casual, rapport-building exchanges.

4.2. Class Engagement and Participation

The shift in ESL classroom engagement and participation was also found to be a prominent theme in the impact of rapport on the development of ESL learners' communicative competence. Whole-class interactions face changes across modalities, as there is unequal participation among students and reduced class dynamics. With the exception of one of the participants, all the participants expressed their inclination towards face-to-face classes in terms of engaging and participating in the language classroom, compared to hybrid and online classes.

Respondent 12 reported, "I have a low participation rate in communicating in online session with other students, as it feels more structured and restricted". The online classroom environment feels more structured for learners, which causes them to reduce participation and engagement in the ESL classroom, affecting the whole-class engagement. This is also supported by Respondent 2, comparing the student's intention to participate in whole-class discussions in physical classrooms as opposed to online classrooms, "During face to face I am more active and energetic as I can speak out most of my ideas and thinking while during online session I tend to become more passive".

The transition to hybrid learning models has redefined full-class dynamics, where the immediacy of interactions typically seen in traditional classrooms is often replaced by a more segmented and less dynamic exchange in virtual environments. In relation to this, Respondent 6 mentioned that "classes feel less interactive, and sometimes it's awkward to have conversations because people are quiet or distracted". Interactions in online and hybrid classes are often less dynamic due to the reduced engagement from students. Research suggests that while hybrid classrooms can integrate diverse learning methods, they frequently struggle to promote inclusive, lively discussions among ESL learners (Ramalingam et al., 2023; Yang & Foley, 2024). This limitation can hinder the development of collective communicative competence (Roinah et al., 2024), as students may feel less inclined to engage in discussions in less personalized online settings.

Comparing face-to-face, hybrid, and online classrooms, Respondent 8 stated, "In face-to-face classes, it's easier to build connections and engage in group activities, while online or hybrid settings sometimes make communication feel more distant". Discussing group work which usually comes naturally and easily for the students in face-to-face classrooms becomes challenging in online classrooms which requires additional settings in the software used or the implementation of third-party software. At the same time, the same respondent also pointed out that "However, the supportive relationships with both the instructor and other students have positively influenced my overall progress" when explaining their communication in the online classroom, which implies that online classrooms can also be engaging with the presence of supportive instructors and classmates.

The nature of full-class discussions has transformed significantly due to the transition to online learning. Alawamleh et al. (2020) discuss the shift in classroom dynamics whereby online learning has hampered robust exchanges typical of in-person classes, thus affecting collaborative learning opportunities. Long et al. (2021) further elaborate by examining how pandemic-

induced limitations on social interactions alter classroom discourse, suggesting that the relational capital that fuels these exchanges has been considerably weakened.

The findings show that class engagement and participation shift noticeably across modalities, shaping how rapport supports ESL learners' communicative competence. Research echoes these experiences, noting that virtual and hybrid environments often struggle to foster inclusive dialogue, ultimately limiting opportunities for developing collective communicative competence.

4.3. Asking and Addressing Questions

Due to the change in modalities, students experience difficulty in asking questions, hesitation to interrupt for asking questions, and uncertainty in expressing themselves in written form. These have a significant impact on the rapport in the language classroom, undeniably also affecting interaction and their development of communicative competence.

Respondent 20 notes the ease of asking questions to the ESL instructor in face-to-face classrooms, "It is easier for me to communicate, or even ask questions to the instructor since most of the classes are face-to-face. I believe that it is also easier for the instructor to see our facial expressions and offers help when needed". Asking questions is deemed easier in physical classrooms as the ESL instructor would be able to see students' facial expression and body language which would cue their need for assistance. At times, learners might want to ask quick questions, which would be easily done in face-to-face classrooms but becomes difficult in online classrooms. This is expressed by Respondent 8, "In online settings, it's harder to casually chat or ask quick questions like in a classroom".

The modalities employed in ESL instruction also influence how learners approach asking and addressing questions. In face-to-face settings, the familiarity and comfort derived from non-verbal exchanges often encourage more proactive inquiry among students (Hamid et al., 2025). However, studies reveal that in online or hybrid contexts, students may hesitate to engage in verbal questioning due to fears of miscommunication or being misunderstood, reflecting a shift in classroom cognitive engagement (Nesprava et al., 2023; Felcida & Deepa, 2025). The challenge of facilitating an environment conducive to questioning within virtual settings compounds the difficulty ESL learners face (Rahman & Razali, 2024; Ying et al., 2021).

However, the shift to questions in written form via chatbox and messages has also raised other concerns. Respondent 12 stated, "For online classes, I feel a bit insecure about my writing skills as I'm required to write a formal question or statement when needed. Sometimes, I tend to keep the question to myself as I found it hard to express my thoughts through a writing form". Students focus too much on expressing themselves formally and accurately when asking questions in written form, hence halting them from actually asking questions in the ESL classroom. Changes in how students ask and address questions within these varying formats also warrant discussion. Recent research indicates that students may be less likely to engage in interaction when they feel distanced from the preferred modalities of learning (Towner et al., 2022).

Undoubtedly, there are learners who thrive in online situations because of the opportunity to ask questions in written form. Respondent 17 said, "When face-to-face, it is hard for me to ask a question as I'm shy, but in online class I can ask anytime," expressing the convenience of asking questions in the chatbox where they can type in the question to be answered by the instructor without disrupting the class. This student noted that the absence of immediate peer

evaluation and the availability of chat functions reduced anxiety. This divergence may be linked to individual personality traits, prior experience with online learning, or higher self-esteem in written proficiency compared to oral production. However, as Peimani and Kamalipour (2021) reported, while some respondents benefited from the immediacy of live online sessions, the lack of immediate feedback on questions could discourage engagement and result in a more passive learning experience.

The findings show that shifts in modality significantly affect how ESL students ask and address questions, which in turn shapes rapport and opportunities for developing communicative competence. In face-to-face classrooms, students feel more comfortable seeking help because instructors can read their non-verbal cues and respond immediately. Overall, the modality influences students' confidence and willingness to ask questions, directly impacting rapport and communicative competence development.

4.4. The Role of Spoken Communication

When there is a shift in modality, the role of spoken communication becomes peripheral, where students rely more on written or asynchronous communication. Students tend to use the chat-box function and prefer sending the teacher messages after class. A few of the students felt uncomfortable to interrupt by asking questions in online classes and felt more comfortable to express their thoughts via chat or written communication.

Respondent 5 mentioned, "For online classes, the way I communicate with my lecturer is just through the chatbox during our online classes or through WhatsApp", With online classrooms, students choose to communicate more via written communication such as the chat function and other social media communications. For many ESL learners, the dynamics of classroom discussions have shifted significantly with the integration of technology, leading to reduced opportunities for spontaneous verbal exchanges that often enhance language skills (Rahman & Razali, 2024; Yang & Foley, 2024). What would normally be a classroom of learners responding and asking questions verbally in whole-class interactions has become less spontaneous, with reduced opportunities for authentic unrehearsed communication in the target language. This is supported by research which indicates that while online tools can facilitate structured communication, they may not replicate the organic nature of spoken discourse found in physical classrooms (Ramalingam et al., 2023).

A review by Singh et al. (2021) also points out how the blend of online instruction and remote learning typically results in less emphasis on verbal interactions, which undermines the development of vital speaking skills among ESL learners. As synchronous communication often lacks the spontaneous exchanges found in face-to-face settings, students may not engage as deeply, resulting in a superficial engagement with the material. However, some respondents, such as Respondent 3, have maintained that written communication is not necessarily a deterrent for developing bonds with fellow classmates. Respondent 3 said, "It can feel closer in face-to-face settings because we interact more directly, but online or hybrid learning allows us to stay connected through group discussions and chats". Students do not necessarily need to be physically present in the classroom to communicate, build rapport, and develop their communicative skills in the English language classroom.

According to the findings of this study, this shift in modality reduces opportunities for spontaneous spoken communication in ESL classrooms, pushing students to rely more on written messages, which alters how rapport and communicative competence develop. Many students

feel hesitant to interrupt or speak in online settings, resulting in fewer authentic verbal exchanges and a decline in natural speaking practice. Research similarly notes that online tools, while useful for structured interaction, cannot fully replicate the spontaneity of face-to-face discourse.

4.5. After-Class Bonding

The pandemic has restricted opportunities for out-of-class interactions, which are vital for reinforcing language skills and fostering rapport among learners. In face-to-face classrooms, students and teachers would often have authentic meaningful conversations either about the course or other matters. These would not only provide learners with the opportunity to practice their English language skills, but they would also open doors for rapport building.

All the participants reported that they had a better social and informal relationship with their instructors and other students when they attended face-to-face classes. Respondent 11 mentioned, “In face-to-face settings, I communicate more casually and frequently with other students, while in online settings, it’s more formal and task-oriented due to the limited interaction opportunities outside class discussions”. Online communication is limited to the academic and class related matters, as the concept of after-class communication is diminished by the “End call” button. After-class interactions are particularly essential for ESL students, as they often rely on immersive experiences to contextualize their learning (Rahman & Razali, 2024).

Respondent 19 also highlighted the important role of after-class communication with teachers, “As I felt more connected with my instructor, it felt easier for me to communicate with her throughout the class and even after class”. This allows for a stronger connection between the teacher and students, especially in developing their English communication skills, and improving their confidence and motivation in using the language. The limited scope for socialisation outside of structured learning environments places additional barriers on language practice, which can detrimentally impact communicative competence and confidence (Hernandez-Ledezma & Janecek, 2025).

Opportunities for informal after-class interaction which are important for rapport and language practice have diminished in online settings, where communication becomes strictly task-focused. Online learning restricts these interactions, reducing valuable spaces for practising English and building rapport.

4.6. Non-Verbal Communication.

Non-verbal communication plays a pivotal role in establishing rapport, an aspect that has been notably diminished in online and hybrid settings. Face-to-face classrooms allowed for all the participants in this study to express themselves more effectively using non-verbal communication.

Respondent 6 felt a stronger connection with the instructor in face-to-face language classes. “In face-to-face classes, I feel more connected with the instructor because we can talk directly, and it’s easier to understand expressions and body language”. Online classes do not offer the same opportunities to learners. Research indicates that the absence of physical cues in digital environments can lead to misunderstandings among ESL students, affecting their learning experience and communicative abilities (Ying et al., 2021). The reliance on verbal communication

alone may not suffice for effective interactions, particularly for ESL learners who depend heavily on contextual clues present in face-to-face interactions (Nesprava et al., 2023). Moreover, studies suggest that while online platforms facilitate certain types of engagement, they often lack the depth of interaction that physical classrooms provide (Ramalingam et al., 2023; Hamid et al., 2025).

Respondent 1 also highlights the limitations in terms of non-verbal communication when it comes to online classroom interactions. “Non-verbal cues, such as body language, can help convey understanding or confusion, which can be harder to interpret in online formats,” said Respondent 1. This lack of non-verbal communication manifests as misunderstandings during interactions, particularly for ESL learners who rely on facial expressions and gestures to grasp contextual nuances (Singh et al., 2021). Though the literal meaning conveyed by teachers or other students may be easily understood, some ESL learners may struggle to make sense of the intended or underlying learning, especially when receiving feedback or comments from others.

These findings suggest that non-verbal communication is crucial for building rapport in ESL classrooms, but it is significantly reduced in online and hybrid settings. Although digital platforms support basic communication, they cannot fully replicate the depth and clarity provided by non-verbal cues in physical classrooms. As a result, the reduced availability of such interactional cues may limit opportunities for spontaneous communication and feedback, which are important for supporting the development of students’ communicative competence in the English language classroom.

5. Conclusion and Implications

Overall, this study highlights the significant shifts that hybrid and online modalities have brought to the ESL learning experience, revealing the complex interplay between teaching formats, socialisation, and student engagement. The transition to online and blended learning during the pandemic has exposed multiple challenges that shape how learners interact, communicate, and develop rapport in language classrooms. Among these challenges, reductions in spontaneous spoken interaction and weakened rapport emerged as the most consistently reported and damaging for learners’ communicative engagement across modalities. Understanding these changes is critical for informing pedagogical strategies that can effectively support communicative competence in diverse learning environments.

The implications for ESL pedagogy in online and hybrid contexts underscore the need for prioritised interventions that directly address the most salient modality-related constraints. Evidence from most participants indicates that reduced spoken output and hesitation to initiate interaction were most pronounced in fully online settings, suggesting an urgent need for low-effort, high-impact instructional adjustments. Immediate pedagogical actions should therefore prioritise structured opportunities for oral interaction, such as breakout rooms, small-group discussions, and task-based speaking activities, which directly target reduced participation and support rapport rebuilding.

Equally important is sustaining personalised support and feedback, which can be compromised in online settings. While reported by several participants rather than the majority, delays in feedback and reduced instructor immediacy were nevertheless linked to disengagement and

uncertainty. Medium-term interventions, such as targeted professional development for instructors on multimodal feedback and online rapport-building strategies, may therefore be necessary to maintain relational quality and learner confidence across hybrid and online modalities.

Beyond instructional strategies, the study also emphasises the social dimension of language learning. Online environments can inadvertently reduce peer interaction, limiting informal practice and collaborative opportunities that are essential for fluency and confidence. To address this challenge, instructors may consider incorporating structured opportunities for informal interaction, such as allocating 10–15 minutes of unstructured speaking time within a 90-minute lesson through breakout-room discussions, rotating speaking partners regularly to broaden peer engagement. Interactional competence may also be supported through assessment practices that recognise participation and interactional fluency, for example by integrating a small percentage of course marks for collaborative speaking tasks or discussion-based activities. At a broader level, curriculum design that intentionally embeds peer collaboration, social learning outcomes, and interactional competence can help mitigate modality-related disengagement and sustain communicative development across face-to-face, hybrid, and online learning environments.

Although this issue was more evident among a subset of participants, its cumulative impact on communicative confidence suggests the need for longer-term curricular attention. Curriculum-level design that intentionally integrates peer collaboration, social learning outcomes, and interactional competence can help mitigate modality-related disengagement over time. By aligning immediate, medium-term, and long-term interventions with the relative strength of evidence, ESL educators can mitigate the challenges posed by online modalities and create environments that continue to support effective language acquisition and communicative competence development.

While the findings offer pedagogical insights relevant to ESL instructors and higher education stakeholders, this study is based on a single qualitative case within one institutional context, which limits the generalisability of the findings. The results are therefore intended to support context-sensitive pedagogical decision-making rather than prescriptive policy formulation.

Future research could strengthen the evidence base through multi-site studies across diverse ESL contexts, as well as mixed-methods designs that combine interview data with classroom observation, survey measures of willingness to communicate, or longitudinal tracking of communicative development. Such approaches would allow for broader validation of the patterns identified in this study and support more robust institutional and curricular decision-making.

Importantly, this study does not seek to directly measure changes in learners' communicative competence as a linguistic construct. Rather, the findings illuminate how shifts between face-to-face, hybrid, and online modalities influence the interactional conditions that support the development of communicative competence, particularly through the role of rapport in shaping learners' willingness to participate, ask questions, and engage in spoken interaction. Participants' accounts primarily reflected changes in their comfort, engagement, and readiness to communicate across modalities, which are widely recognised in second language research as enabling conditions for communicative practice and competence development. In this sense, the study contributes by highlighting how modality-related changes in rapport and interactional dynamics may expand or constrain opportunities for communicative engagement within ESL classrooms. Understanding these interactional mechanisms provides a valuable perspective for

educators seeking to create learning environments that sustain meaningful communication and support the longer-term development of communicative competence.

It should be noted that the interviews were conducted in 2024, while the instructional modality shifts experienced by participants occurred between 2020 and 2022 during the COVID-19 pandemic. As a result, participants were asked to recall and compare learning experiences that occurred several years earlier. Retrospective accounts may be subject to memory distortions, including recall bias or the reconstruction of past experiences in ways that align with present perspectives. Although the interviews focused on salient learning experiences and interactional moments that participants could recall clearly, the temporal gap may influence the completeness or accuracy of some accounts. Future research may benefit from collecting data closer to the time of instructional changes or using longitudinal designs to capture evolving learner experiences.

Acknowledgments

The authors would like to acknowledge Universiti Putra Malaysia for funding this research via GP-IPM/9770500.

Declaration

The authors declare that artificial intelligence (AI) tools were used during the preparation of this manuscript to assist with language editing, organisation of ideas, and improvement of clarity and academic expression. All content was critically reviewed, revised, and validated by the authors, and the responsibility for the accuracy, originality, and integrity of the work rests entirely with the authors.

References

- Agostinelli, M. D. (2019). *From distance education to online education: A review of the literature* (Graduate student thesis). Graduate Student Theses, Dissertations, & Professional Papers.
- Alaskar, H. (2023). The role of online learning in enhancing the performance of introverted female Saudi students in translation. *Saudi Journal of Language Studies*, 3(3), 158–182. <https://doi.org/10.1108/sjls-12-2022-0092>
- Alawamleh, M., Al-Twait, L., & Al-Saht, G. (2020). The effect of online learning on communication between instructors and students during COVID-19 pandemic. *Asian Education and Development Studies*, 11(2), 380–400. <https://doi.org/10.1108/AEDS-06-2020-0131>
- Almendingen, K., Morseth, M. S., Gjølstad, E., Brevik, A., & Tørris, C. (2022). Student experiences with online teaching during COVID-19: A qualitative study of higher education learning environments. *Education Sciences*, 12(3), 1–14. <https://doi.org/10.3390/educsci12030189>
- Ash'ari, N., Azmi, F., Navarathnaraja, K., Tan, W., Yunus, M., Hashim, H., & Syafril, S. (2025). Student perceptions of metacognitive strategies in hybrid ESL classrooms: Advancing self-regulated learning for future-ready education. *International Journal of Future Ready Learning and Education*, 39(1), 38–47. <https://doi.org/10.37934/frle.39.1.3847>
- Belda-Medina, J. (2021). Enhancing multimodal interaction and communicative competence through task-based language teaching in synchronous computer-mediated communication. *Education Sciences*, 11(11), 723. <https://doi.org/10.3390/educsci11110723>
- Cho, M.-H., & Cho, Y. (2016). Online Instructors' Use of Scaffolding Strategies to Promote Interactions: A Scale Development Study. *The International Review of Research in Open and Distributed Learning*, 17(6). <https://doi.org/10.19173/irrodl.v17i6.2816>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- Denker, K., Knight, K., Carroll, R., Bradley, K., Bonine, P., Lauck, S., ... Storr, M. (2022). Assessing student mindset, interest, participation, and rapport in the post-pandemic public speaking classroom: Effects of modality change and communication growth mindset. *Journal of Communication Pedagogy*, 6, 178–194.
- Fehérvári, A., Major, E., & Széll, K. (2022). Student engagement and experiences of online learning during the COVID-19 pandemic in higher education. *Education and Information Technologies*, 27, 7731–7748. <https://doi.org/10.1007/s10639-022-10925-6>
- Felcida, F., & Deepa, P. (2025). An experimental study on developing competency of students among L2 learners through hybrid teaching and learning in selected engineering colleges. *Mitteilungen Klosterneuburg*. <https://doi.org/10.61586/s2she>

Frisby, B., & Martin, M. (2010). Instructor–student and student–student rapport in the classroom. *Communication Education*, 59(2), 146–164.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>

Halliday, M. A. K. (1978). *Language as a social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Halliday, M. A. K. (2004). Three aspects of children’s language development. In J. Webster (Ed.), *The language of early childhood* (pp. 308–326). Continuum.

Hamid, R., Noni, N., & Sunra, L. (2025). The challenge of hybrid learning in teaching English for business. *Klasikal Journal of Education, Language Teaching and Science*, 7(1), 381–389. <https://doi.org/10.52208/klasikal.v7i1.1312>

Hernandez-Ledezma, L. J., & Janecek, U. (2024). AI-guided professional development: A Vygotskian approach for teachers in multilingual classrooms. In S. C. Kitchens & J. Calhoun (Eds.), *Harnessing AI’s potential to support student success and teaching excellence* (pp. 245–276). IGI Global. <https://doi.org/10.4018/979-8-3373-1122-7.ch009>

Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.

Junn, H. (2023). L2 communicative competence via synchronous computer-mediated communication as an alternative to formal classrooms. *Innovation in Language Learning and Teaching*, 17(1), 15–31. <https://doi.org/10.1080/17501229.2021.1895802>

Liu, J., Wyver, S., & Chutiyami, M. (2022). Impacts of COVID-19 restrictions on young children’s outdoor activity: A systematic review. *Children*, 9(1564). <https://doi.org/10.3390/children9101564>

Long, E., Patterson, S., Maxwell, K., Blake, C., Pérez, R., Lewis, R., ... Mitchell, K. (2022). COVID-19 pandemic and its impact on social relationships and health. *Journal of Epidemiology & Community Health*, 76(2), 128–132. <https://doi.org/10.1136/jech-2021-216690>

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Nesprava, M., Zarishniak, I., Koropetska, O., Kurinna, A., Syvash, S., & Sushchenko, L. (2023). Formation of communicative competencies in higher education in distance learning. *Revista Românească pentru Educație Multidimensională*, 15(1), 247–260. <https://doi.org/10.18662/rrem/15.1/695>

Ong, S. G. T., & Quek, G. C. L. (2023). Enhancing teacher-student interactions and student online engagement in an online learning environment. *Learning environments research*, 1–27. Advance online publication. <https://doi.org/10.1007/s10984-022-09447-5>

Peimani, N., & Kamalipour, H. (2021). Online education in the post-COVID-19 era: Students' perception and learning experience. *Education Sciences*, 11(10), 633. <https://doi.org/10.3390/educsci11100633>

Rahman, S., & Razali, A. (2024). Adapting to change: Facing challenges in developing ESL communicative competence online. *Pertanika Journal of Social Sciences and Humanities*, 32(1), 1–16. <https://doi.org/10.47836/pjssh.32.1.01>

Ramalingam, S., Yunus, M., & Hashim, H. (2023). Exploring ESL students' insights on learner-to-learner interaction in blended environments. *Global Conference on Business and Social Sciences Proceedings*, 15(1), 114. [https://doi.org/10.35609/gcbssproceeding.2023.1\(114\)](https://doi.org/10.35609/gcbssproceeding.2023.1(114))

Roinah, R., Sofiani, I., Susanti, T., Sumarni, T., & Mahsuri, M. (2024). Blended learning to enhance productive skills. *International Journal of Innovative Research in Multidisciplinary Education*, 3(10). <https://doi.org/10.58806/ijirme.2024.v3i10n04>

Saltz, J., & Heckman, R. (2020). Using Structured Pair Activities in a Distributed Online Breakout Room. *Online Learning*, 24(1). <https://doi.org/10.24059/olj.v24i1.1632>

Singh, J., Steele, K., & Singh, L. (2021). Combining the best of online and face-to-face learning: Hybrid and blended learning approaches. *Journal of Educational Technology Systems*, 50(2), 140–171. <https://doi.org/10.1177/00472395211047865>

Smallwood, A. M. K., & Brunner, B. R. (2017). Engaged learning through online collaborative public relations projects across universities. *Journalism & Mass Communication Educator*, 72(4), 442–460. <https://doi.org/10.1177/1077695816686440>

Taskiran, A. (2023). Language learning goes mobile: Online intercultural exchange during COVID-19 pandemic. *MEXTESOL Journal*, 47(2), 1–14.

Towner, E., Tomova, L., Ladensack, D., Chu, K., & Callaghan, B. (2022). Virtual social interaction and loneliness among emerging adults during COVID-19. *Current Research in Ecological and Social Psychology*, 3, 100058. <https://doi.org/10.1016/j.cresp.2022.100058>

Yang, H., & Foley, J. (2024). Translanguaging as interactional resources in EFL student-led group tasks. *Journal of English Language*, 5(1), 33–54. <https://doi.org/10.62819/jel.2024.145>

Ying, Y., Siang, W., & Mohamad, M. (2021). Challenges of English learning and the use of social media/video conferencing during COVID-19. *Creative Education*, 12(7), 1503–1516. <https://doi.org/10.4236/ce.2021.127115>

Appendices

Appendix A: Interview Protocol

Interview Question	Connection to key constructs
a) Do you feel comfortable communicating in English with the instructor in the language classroom? Why?	Gaining an understanding on the general comfort that the participants have for communicating in English
b) Do you feel comfortable communicating in English with other students in the language classroom? Why?	Gaining an understanding on the general comfort that the participants have for communicating in English
c) How do you feel about your relationship with the instructor? Do you feel it has changed depending on the mode of learning (i.e.: online, hybrid, face-to-face)?	Connected to rapport, and how it is related to the change in modality
d) How do you feel about your relationship with other students? Do you feel it has changed depending on the mode of learning (i.e.: online, hybrid, face-to-face)?	Connected to rapport, and how it is related to the change in modality
e) Do you feel that the way you communicate with your instructor has changed depending to the mode of learning (i.e.: online, hybrid, face-to-face)? How? Why?	Connected to communicative competence, and how it is related to the change in modality
f) Do you feel that the way you communicate with other students has changed depending to the mode of learning (i.e.: online, hybrid, face-to-face)? How? Why?	Connected to communicative competence, and how it is related to the change in modality
g) Do you feel that your relationship with the instructor impacts your development of communicative competence? How?	Connected to rapport, and how it is related to the development of communicative competence
h) Do you feel that your relationship with other students impacts your development of communicative competence? How?	Connected to rapport, and how it is related to the development of communicative competence
i) Do you feel that your development of communicative competence has changed depending to the mode of learning (i.e.: online, hybrid, face-to-face)? Would you say it has been	Connected to rapport, and how it is related to the development of communicative competence, specifically in

impacted by your relationship with the instructor or other students?	different modes (online, hybrid and face-to-face)

Appendix B: Example from Codebook

Theme: Class Engagement and Participation Theme Definition: Class Engagement and Participation refers to students' level of involvement, interaction, and communicative contribution during learning activities across different instructional modalities. This theme captures how face-to-face, online, and hybrid learning environments shape students' willingness to participate, their interaction with peers, and opportunities for spontaneous language use.	
Code	Excerpts
Active Verbal Participation	“Yes. During face to face I am more active and energetic as I can speak out most of my ideas and thinking while during online session I tend to become more passive.” (R2) “Yes. In face-to-face classes, I talk more naturally.” (R6)
Engaging Interaction with Peers	“In face-to-face classes, my friends and I can talk freely and casually. We can ask questions to each other directly even during the ongoing lecture or in between breaks.” (R12) “In face-to-face classes, it's easier to build connections and engage in group activities.” (R8)
Restricted Communication in Online Settings	“I have a low participation rate in communicating in online session with other students, as it feels more structured and restricted.” (R12) “Online classes feel less interactive, and sometimes it's awkward to have conversations because people are quiet or distracted.” (R6)

Managerial Governance and Performance Culture: A Case Study of Early-Career Academics' Experiences in a Malaysian Research University

ABSTRACT

Authors: Nordahlia Umar Baki^{1*}
Khairunnisa Mohd Pauzi² & Shamuni
Kunjiapu^{3,4}

Affiliation: ^{1,2}Faculty of Educational
Studies, Universiti Putra Malaysia,
43400 Serdang, Selangor, Malaysia.

³Asia-Europe Institute, Universiti
Malaya.

⁴Department of Business
Management, New Era Universiti
College.

*Corresponding author:

nordahlia@upm.edu.my

Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0
International License

APA citation for this article:

Early-career academics (ECAs) are often the most affected by governance reforms and performance pressures as they try to establish themselves within increasingly managerial university systems. This qualitative case study draws on interviews with six ECAs to explore how they understand and respond to governance practices and key performance indicator (KPI) requirements in a Malaysian research university. The participants comprise an equal number of permanent and contract academic staff, each with less than five years of academic experience. Grounded in Job Demands and Resources (JD–R) theory, the study examines how institutional expectations shape ECAs' daily academic work and influence their early career development. Ethical principles of informed consent, voluntary participation, and confidentiality were observed throughout the research process. The study focused on how governance and KPI systems influence participants' workloads, priorities, and working lives. Three main themes were identified: structuring of daily academic tasks; perceived realism and fairness of expectations; and governance structures as enablers or constraints. Overall, participants described KPIs as a major source of pressure that intensified workload and blurred work–life boundaries, particularly for contract staff and newly returned mothers. Uneven access to mentoring, guidance, and institutional support further widened these pressures, pointing to the need for more transparent, supportive, and career-sensitive governance practices.

Keywords: Early-Career Academics (Ecas), KPI Systems, Managerial Governance, Job Demands–Resources Theory, Malaysian Research Universities.

1. Introduction

The governance of higher education institutions (HEIs) has undergone a profound transformation over the last three decades, characterized by the rise of the "entrepreneurial university" (Slaughter & Rhoades, 2012; Abreu & Grivenich, 2024). This shift has redirected universities from traditional collegial models of self-governance toward corporate-style managerialism, emphasizing efficiency, accountability, and measurable performance outcomes (Lynch & Ivancheva, 2015). Nevertheless, the concept of the entrepreneurial university remains insufficiently defined, primarily due to the institution's inherently complex and adaptive nature. Universities pursue multiple missions, such as advancing research, providing education, encouraging innovation, and engaging with communities, often through different organisational units that have distinct objectives (Abreu & Grinevich, 2024). These varied activities respond to dynamic external pressures; however, their outcomes are seldom examined holistically within a comprehensive analytical framework. Within this evolving context, a performance-oriented culture has become the common practice, where academic value is increasingly evaluated through key performance indicators (KPIs), research funding achievements, and high-impact publications.

Shifts in governance and performance culture have reshaped institutional priorities and, in many contexts, restructured academic employment practices. ECAs are often particularly vulnerable to these changes. Unlike their tenured counterparts who entered the academy during eras of relative stability, today's ECAs experience uncertain career pathways defined by hyper-competition and employment insecurity (Spina et al., 2020; Burton & Bowman, 2022). This intersection of managerial governance and job precarity creates a state of 'cruel optimism' for junior academics (Bone, 2021), compelling them to perform as 'academic superheroes' meeting extreme metrics (high KPIs, impact publications) driven by the false hope of securing increasingly scarce permanent positions (McKenzie, 2018). These affective pressures are not incidental outcomes but are structurally produced through governance and KPI systems that reward visibility, competitiveness, and constant productivity, while offering limited protection or career security for those at the start of their academic careers. These forces ultimately threaten their professional development and sustainable academic work.

In the Malaysian context, RUs play an important role in this transition, leveraging rigorous Performance Management Systems (PMS) as a strategic tool to enhance their pursuit of global academic standards. Under these frameworks, performance is evaluated across a broad set of expectations, including teaching, supervision, research, publication, innovation, and professional service. The scope also extends to community engagement, industry collaboration, student development, and other contributions beyond formal academic duties, creating a comprehensive and demanding evaluation culture. For ECAs, these multi-dimensional PMS may function not only as mechanisms of performance management but also as systems of monitoring and control (Franco-Santos & Doherty, 2017; Harford, 2021). This environment often manifests as an "audit culture," where the intrinsic value of teaching and scholarly integrity can become overshadowed by the demand for measurable outputs (Power, 1997; Strathern, 2000). As a result, ECAs may feel pressured to prioritise quantifiable outcomes over developmental academic activities, reinforcing a system of performativity that encourages strategic effort to meet institutional accountability requirements (Harford, 2021).

Despite the growing body of literature on the general effects of these structural changes (e.g.,

Bone, 2021; Franco-Santos & Doherty., 2017; Teichler, 2019), there remains a pressing need for in-depth study of how these governance structures affect the daily realities of those at the start of their careers, particularly in research-intensive environments (Hollywood et al., 2020; Lee et al., 2021). Most existing studies focus on quantitative metrics of productivity; however, the human cost and the psychological negotiation of these demands require a more micro-level approach. This study contributes to critical higher education in three ways. First, it provides a micro-level analysis of how governance and KPI regimes shape the everyday organisation of work among ECAs, an area underexplored in predominantly macro-level and policy-focused studies. Second, by applying JD–R theory, it offers an analytically grounded explanation of how performance governance operates simultaneously as a source of work intensification and uneven institutional support. Third, the study contributes empirically by foregrounding the experiences of ECAs in a Malaysian research university, extending predominantly global north–centric accounts of managerialism and audit culture.

To ensure conceptual clarity and consistency, key terms used in this study are operationally defined.

- **ECAs** refer to academic staff with fewer than five years of cumulative academic employment, reflecting the formative stage of academic career development.
- **KPIs/ Key Intangible Performance Indicators (KIPs)** are defined as formal performance metrics used by the institution to evaluate academic work across research, teaching, service, and other contributions.
- **Workload** is understood as the totality of teaching, research, administrative, and service tasks required of ECAs, including both formally recognised and informal or invisible work.
- **Job resources** refer to institutional, social, and material supports (e.g., mentoring, autonomy, guidance, funding) that help ECAs meet performance demands and sustain motivation.

These operational definitions informed the development of the interview protocol (Appendix A), which was derived directly from the study's research question: *How do governance and KPI systems shape early-career academics' daily academic work, workload, and priorities?*

2. Literature Review

2.1. Managerial Governance and Its Impact on ECAs

Managerial governance refers to the application of New Public Management (NPM) principle to higher education institutions, characterised by hierarchical control, performance measurement and strict audit mechanism to ensure strategic accountability (Deem et al., 2007). As the academic landscape becomes increasingly entrepreneurial-driven, academics must navigate an expanding set of performance indicators that now evaluate not only teaching and research but also societal impact. Under this system, academic work is no longer understood as a singular vocation grounded in professional autonomy, but fragmented into a bundle of highly specified and measurable services subject to continuous evaluation. In Malaysian universities, this manifests as an increased reliance on short-term contract positions and a strong emphasis on performance-driven KPIs (Wan et al., 2022), developments often associated with growing financial pressures and intensifying competition in the higher

education sector. While these arrangements may provide institutions with greater organisational flexibility, they can also introduce challenges related to academics' well-being, incentives for scholarly innovation, and sustainable workload management (Li et al., 2025).

The first five years of employment are widely recognised as a critical turning point for ECAs, serving as a fundamental foundation for securing future academic rewards and overall career success (Henkel, 2004; Laudel & Gläser, 2008). Conversely, difficulties encountered during this period are strongly associated with eventual job dissatisfaction particularly when governance structures are perceived as unsupportive or misaligned with their interests. The core pressure stems from the expectation that ECAs must excel simultaneously across a complex, multi-component role. ECAs grapple with the pressure inherent in the competitive and often politicised environments of research-intensive universities (Zhuang & Lin, 2025). This complex mandate immediately generates an unsustainable workload. Past literature highlights that the simultaneous pursuit of these diverse and often conflicting missions inherently generates stress and reduces academics' control over time (Franco-Santos & Doherty, 2017; Li et al., 2025). For ECAs, this pressure is magnified because, unlike tenured colleagues, they have limited institutional power to negotiate their commitments. ECAs often carry disproportionately heavy administrative loads (Spina et al., 2020), thereby severely reducing the time available for sustained, high-impact research, which remains the primary metric for career advancement.

In the contexts of Malaysian universities, this imbalance is acute, as national pressures for high-ranking publications often clash with substantial teaching, supervision, and community and professional service demands (Wan et al., 2022). This situation forces a difficult prioritization dilemma: ECAs compelled to prioritize easily measurable research outputs essential for career progression over less-auditable responsibilities (McKenzie, 2018; Harford, 2021), thereby fundamentally reshaping their daily academic priorities. As a result, work intensification becomes normalised, and boundaries between professional and personal life become increasingly blurred. According to Li et al. (2025), ECAs frequently report sacrificing personal time, family obligations, and leisure to manage their exhaustive list of duties due to a culture of compulsory overwork.

2.2. Job Demands-Resources Theory

The Job Demands-Resources (JD-R) theory was designed to be universally applicable to explain the interplay of diverse elements that affect job performance (Bakker & Demerouti, 2007; Demerouti et al., 2001). Specifically, it explicitly explains how excessive job demands such as high workloads, time pressure, and performance monitoring interact with available job resources such as institutional support, autonomy, and feedback to shape employee performance, motivation, and well-being. When applied to this study, the theory offers a valuable framework for understanding how governance and KPI systems shape the work, workload, and priorities of ECAs. In this study, ECAs face rising performance demands under managerial governance while simultaneously losing the traditional autonomy that once protected academic well-being.

JD-R theory posits that every job can be understood through two fundamental categories: job demands and job resources. Job demands refer to organizational aspects that require continuous effort, leading to energy depletion and psychological costs. It may result in stress and burnout

when they exceed an individual's capacity (Bakker & Demerouti, 2007; Bakker et al., 2023). In the “entrepreneurial” research universities, managerial governance and its accompanying KPI systems act as a powerful mechanism for intensifying job demands on ECAs, directly shaping their daily work, workload, and priorities. These demands manifest across multiple dimensions, *i.e.*, quantitative demand, emotional demand, and cognitive demand. Quantitative demands arise from the pressure to simultaneously fulfil the multicomponent academic role of teaching, research, service and networking, often resulting in excessive working hours and unsustainable workloads (Li et al., 2025). Emotional demands emerge when ECAs are expected to perform as “academic superhero” to comply with institutional performance evaluations resulting in emotional labor and psychological strain (Naidoo-Chetty & Du Plessis, 2021). Cognitive demands are evident in the continuous requirement to make strategic prioritisation decisions under KPI pressure, particularly the need to privilege easily auditable performance indicators over deep, developmental scholarly work. This enforced trade-off not only intensifies mental load but also risks compromising ECAs' long-term scholarly development (Lynch, 2015; Strathern, 2000). Collectively, these escalating job demands provide a robust explanatory framework for understanding the heightened stress, work intensification, and performance tensions experienced by ECAs in KPI-driven governance environments (Li et al., 2025).

Conversely, job resources refer to the physical, social, and organisational aspects of work that facilitate goal attainment and personal growth, while enabling performance, buffering the strain induced by job demands, and supporting the fulfilment of fundamental psychological needs (Bakker & Demerouti, 2007; Bakker et al., 2014). Importantly, the JD-R theory emphasises that an individual's perception of a demands (as a threat or a challenge) is contingent upon the availability of these resources and their coping capacity. Based on the literature review, managerial governance emerges as a dual mechanism that not only intensifies job demands but also actively erodes or restricts the job resources traditionally available to academics (*e.g.*, autonomy and social support) (Ani Marlia et al., 2025; Huang, 2024; Zhuang & Lin, 2025). Highly specified KPIs and centralised control structures operate to constrain academic freedom, consequently weakening the autonomy required for ECAs to independently organise their work schedules (Deem et al., 2007). The inability to negotiate workload or challenge unrealistic targets (KPIs) exacerbates strain, as ECAs are effectively denied the resource of control over their work environment (Franco-Santos & Doherty, 2017). Furthermore, hyper-competitive, performance-based systems, which create rivalries for scarce rewards, might undermine collegial trust and cooperation, thereby eroding social resources (Lynch, 2015). Under the JD–R model, such patterns suggest a dual process in which escalating demands and diminishing resources jointly contribute to heightened strain, reduced motivation, and intensified work effort among ECAs operating within KPI-driven governance contexts. This theoretical reasoning provides the conceptual basis for examining how governance structures shape ECAs' daily work experiences in the present study. The resulting conceptual framework is illustrated in Figure 1.

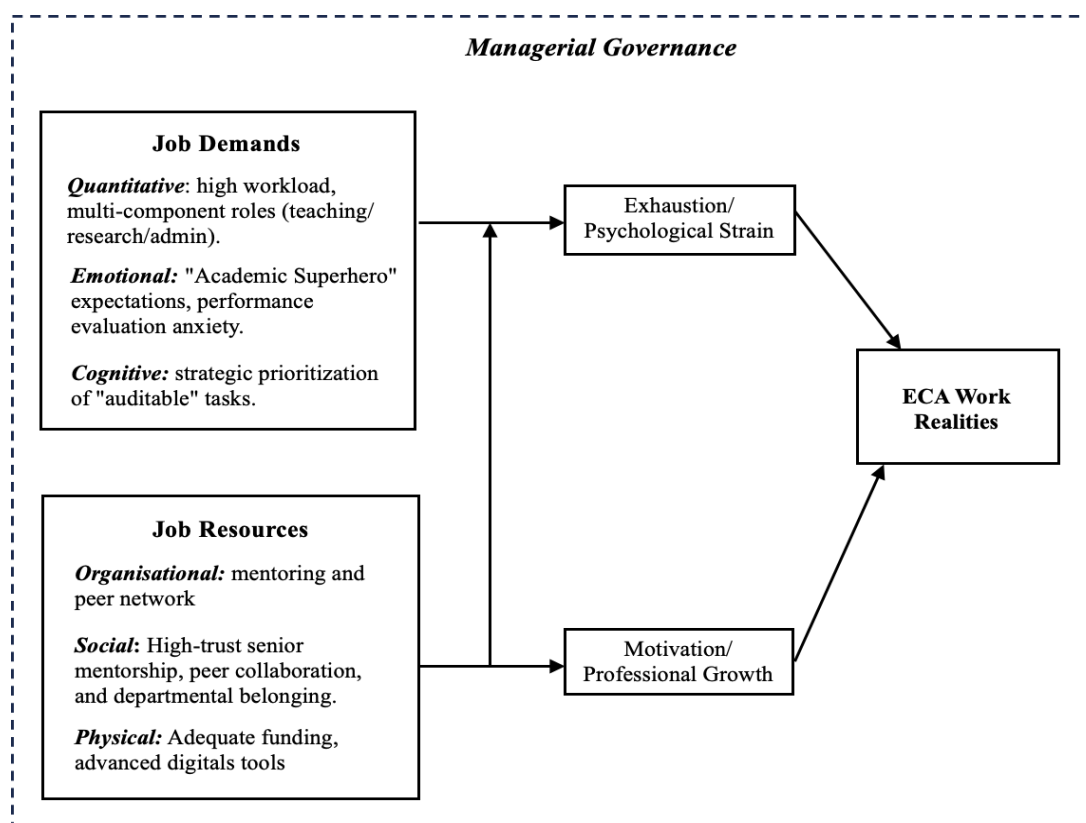


Figure 1. Conceptual Framework Linking Managerial Governance, JD-R and ECAs' Work Experiences

3. Methodology

This study employed a qualitative research design to explore the governance and performance culture experienced by ECAs within a Malaysian research university. The study was conducted at a public, research-intensive university in Malaysia, classified as one of the country's Research Universities (RUs) under the national higher education framework. Established more than five decades ago, the institution is a large, multidisciplinary university with a strong strategic emphasis on research productivity, international publication performance, and performance-based academic evaluation systems. Consistent with qualitative case study methodology, the institution constitutes a bounded case within which ECAs' experiences were examined in relation to governance and KPI practices. These institutional characteristics situate the case within Malaysia's highly competitive research-intensive university sector, where formalised governance and performance management systems strongly shape academic work. To protect institutional anonymity, specific identifying details are withheld. Eligible participants were identified across several faculties based on the study's inclusion criteria (academic staff in either contract or permanent positions with fewer than five years of service). From this pool, twelve early-career academics were contacted through university email directories and invited to participate. Six individuals agreed within the study timeframe, forming the final sample. The participants included both contract-based and permanent academic staff.

This sample size aligns with the exploratory nature of qualitative inquiry, where depth of insight and contextual understanding are prioritised over breadth of representation. In qualitative

interview research, relatively small samples can generate rich analytical insights when participants share comparable professional contexts and experiences. During the coding process, the researchers observed that the core themes relating to workload structuring, KPI pressures, and governance expectations began to recur across interviews, and by the later interviews no substantially new thematic categories were emerging. This approach aligns with qualitative research guidance suggesting that thematic saturation can often be achieved within relatively small samples in focused interview studies (Guest et al., 2006). Evidence of thematic saturation was monitored through iterative code accumulation across interviews, as summarised in Appendix B.

The inclusion of participants with less than five years of experience was justified on the grounds that ECAs are typically in the formative stage of their academic careers, navigating institutional governance structures, adapting to performance expectations, and negotiating career identity within the pressures of managerialism. Including both contract and permanent staff allowed for the comparison of differing employment contexts, offering a more nuanced understanding of how governance and KPI mechanisms influence academic work, professional development, and institutional engagement. Participants were selected based on the following inclusion criteria:

- Individuals currently employed within one Malaysian RU as academic staff.
- Those with less than five years of cumulative academic work experience.
- Academics occupying either contract or permanent positions.
- Willingness to provide informed consent for participation and data use

The primary method of data collection was in-depth, semi-structured interviews ranging from 60 to 90 minutes. The interview recordings were transcribed verbatim and analysed using ATLAS.ti version 23. A reflexive thematic analysis approach was adopted to identify patterns across participants' experiences. The analysis proceeded in several stages. First, transcripts were read repeatedly to familiarise the researchers with the data. Initial open coding was then conducted to identify meaningful segments related to governance perceptions, workload structuring, and KPI expectations. In the second stage, related codes were grouped into broader categories through iterative comparison across transcripts. Finally, these categories were refined into higher-level themes that captured recurring patterns across participants' experiences. To enhance analytical rigour, coding decisions and theme development were reviewed by two researchers through peer checking, and differences in interpretation were discussed until consensus was reached. As a reflexive thematic analysis approach was adopted, inter-coder reliability statistics were not calculated; instead, analytic credibility was supported through iterative discussion, reflexive memoing, and consensus-building among the research team. An excerpt of the coding framework illustrating themes, selected codes and definitions is shown in Table 2.

The interview protocol focused on governance perceptions, the impact of performance metrics on daily academic work, and their emotional and motivational consequences, as outlined in Appendix A. While semi-structured interviews constituted the primary analytic data source, institutional documents (e.g., KPI frameworks, promotion guidelines, and strategic plans), observational notes, and researcher reflective notes were used to support contextual understanding, inform interpretation, and enhance reflexivity, rather than as independent datasets for formal thematic analysis. Accordingly, the findings reported in this paper are

grounded in interview data, with supplementary materials used to enhance analytical sensitivity and credibility.

Next, peer checking was conducted by engaging two researchers to review coding decisions and thematic interpretations, helping to minimize individual bias and improve analytical rigor. These measures supported the credibility, dependability, and confirmability of the study's findings, ensuring that interpretations remained grounded in participants' lived experiences and institutional realities. Throughout the research process, the researchers engaged in reflexive practice to remain aware of how prior familiarity with higher education governance and performance systems could shape interpretation. Reflexive memos were used to bracket assumptions, interrogate moments of analytic resonance, and ensure that themes remained grounded in participants' accounts rather than researcher expectations.

The study was conducted in accordance with the university's internal code of conduct for research involving human participants and internationally recognised ethical principles governing qualitative research, including the principles of respect for persons, beneficence, and justice articulated in the Belmont Report (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979) and the British Educational Research Association Ethical Guidelines for Educational Research (British Educational Research Association, 2018). Several safeguards were implemented to protect participants throughout the study. First, all participants received an explanatory statement outlining the purpose of the research, the voluntary nature of their participation, their right to withdraw at any time without penalty, and the intended use of the data. Second, verbal informed consent was obtained prior to the commencement of each recorded interview. Third, pseudonyms (P1–P6) were assigned to all participants, and identifying information such as department names and specific responsibilities was removed from the transcripts to ensure confidentiality. Fourth, interviews were conducted either virtually or in neutral off-campus locations to allow participants to speak freely without concern for institutional oversight. Finally, all interview recordings and transcripts were stored on a restricted cloud drive accessible only to the principal researcher, and anonymised research data will be securely retained for five years in accordance with recognised research data management practices.

4. Findings and Discussion

The study involved six ECAs who were assigned pseudonyms to protect their identities (P1–P6), as shown in Table 1. The sample comprised four female and two male participants, reflecting variation in gender composition. Three participants held permanent academic positions (P1, P2, P3), while the remaining three were employed on contract basis (P4, P5, P6). Their academic working experience ranged from 2 to 3.5 years, consistent with the study's focus on staff with less than five years in academia. This distribution enabled the exploration of how employment status (permanent versus contract) and early career stage intersect with experiences of governance and performance measurement in higher education.

Table 1. Participant Demographic

Participant (Pseudonym)	Gender	Employment Status	Working Experience (years)
P1 (F)	Female	Permanent	3.5
P2 (N)	Female	Permanent	2.5
P3 (Z)	Male	Permanent	3
P4 (S)	Female	Contract	3
P5 (D)	Female	Contract	2
P6 (R)	Male	Contract	4

The findings and discussion are structured around the central research question: *How do governance and KPI systems shape ECAs' daily academic work, workload, and priorities?* Three main themes and eight sub-themes emerged from the analysis and are presented in Table 2, followed by detailed discussion in the sections that follow.

Table 2. Governance and KPI system in shaping ECAs' work

Theme	Sub-theme	Working definition	Exemplar codes
Structuring of daily academic tasks	KPI-driven task prioritisation	Ways in which formal KPIs determine which activities ECAs prioritise or postpone in their daily schedules.	“chasing publications”, “grant first, teaching later”, “activities generating more income”, “deadline-driven tasks”
	Balancing teaching, research, and professional service	How ECAs distribute time and effort across teaching, research, and service in response to institutional performance expectations.	“overloaded hours with the rest of KPI”, “night-time research”, “take leave to focus on research”, “juggling multiple roles”, “split focus”, “working on weekend”
	Administrative and reporting burden	Extra workload related to documenting, tracking, and reporting KPIs and governance requirements.	“online KPI system data entry”, “filling forms”, “evidence uploading”, “time spent on reports”, “paperwork burden”, “delaying process”, “various system which not merging to one another”

Perceived realism and fairness of expectations	Realistic versus unrealistic KPI targets	ECAs' judgements about whether KPI targets are attainable given their career stage, and resources.	"targets too high", "unrealistic for new lecturer", "benchmarks for senior staff applied to juniors" "insufficient mentoring", "same KPI for all DS13 although different year of service", "no consideration of learning curve"
	Differences by employment status and gender	How employment status shapes exposure to and experience of KPI pressures.	"KPI tied to contract renewal", "job insecurity", "extra pressure on contract staff", "KPI set (in contract) sometimes no aligned in the system", "inequality of risk", "newly working mom received less support"
Governance structures as enablers or constraints	Supportive versus restrictive management practices	Ways in which faculty or university management either facilitates or obstructs ECAs' work under KPI regimes.	"supportive HOD", "rigid rules", "top-down decisions", "flexible workload allocation", "pressure from management", "too much bureaucracy",
	Transparency and communication of expectations	Clarity, visibility, and communication of policies, KPIs, and performance criteria.	"unclear criteria", "late communication", "policy confusion", "orientation/briefing sessions", "learning by trial and error"
	Autonomy in workload negotiation	ECAs' perceived ability to negotiate tasks, redistribute workload, or resist unreasonable demands.	"no say in workload", "negotiating teaching hours", "fully autonomy in class", "pushing back on tasks", "limited voice", "accepting decisions"

Theme 1: Structuring of daily academic tasks

Analysis of the data revealed that the theme structuring of daily academic tasks comprised several interrelated sub-themes, namely: (a) KPI-driven task prioritisation; (b) balancing teaching, research and professional services; and (c) administrative and reporting burden.

a. KPI-driven task prioritisation

Participants consistently described how KPIs shaped the organisation of their daily work, often redirecting their efforts towards activities that were formally measurable and reportable (Javed

& Alenezi, 2023). Work associated with quantifiable outputs particularly publications, grants, and other research-related indicators was routinely prioritised over teaching preparation, student support, and service responsibilities. Several participants explicitly characterised their experiences as “chasing publications” and adopting a “grant first, teaching later” orientation, suggesting that research and income-generating activities, such as organising conferences and seminars, increasingly dominated their academic schedules. As P5 expressed:

“Sometimes I feel like an event organiser more than an academic, so much energy goes into planning, logistics, and reporting for activities that are mainly there to generate income and KPIs.” (P5)

This perception was echoed by P2, who described how substantial time and effort were diverted from core academic responsibilities towards revenue-oriented tasks that were viewed as institutionally prioritised, albeit with unclear benefits for staff development:

“The management has a lot on their agenda that we don’t fully see or understand in terms of what it brings to ECAs. What I do see is that they focus on generating revenue or income, and we end up organising a lot of conferences to generate revenue, which are not even part of our main KPIs.” (P2)

For P6, these conditions resulted in an increasingly instrumental approach to work organisation, described as “working backwards from the KPI list,” whereby institutional metrics and appraisal timelines became the primary reference points for planning weekly and semester activities. This produced a predominantly deadline-driven workflow, in which work associated with imminent KPI-related cut-offs took precedence. Consequently, several participants acknowledged engaging in strategic task management by deferring or minimising responsibilities that were less visible within formal performance systems such as mentoring, collegial support, industry attachment and community development in order to preserve time and energy for activities that were institutionally recognised, rewarded and visible to the management. These experiences were broadly consistent with institutional documents that emphasise publication output and performance-based evaluation criteria.

b. Balancing teaching, research, and professional services

Participants described significant challenges in balancing teaching, research, and professional service, including industry attachments and community development, under institutional performance expectations. Many reported that official workloads did not capture the invisible time required to meet KPIs, creating a sense of constant juggling across competing responsibilities. This often meant adopting a “long hours” culture, being available 24/7, working nights, weekends, holidays, or taking leave just to complete research, plan lessons, or manage emails (Bozzon et al., 2017). Several participants described “night-time research” (P3) and “working on weekends” (P4, P6) as a common strategy, while others had to “take leave to focus on research” (P1, P2, P4) because regular working hours were consumed by teaching, meetings, and service tasks. As Bozzon et al. (2017) note, “the boundaries between work and other life spheres seem to be weak, in an ambiguous exchange between work vocation and

precariousness,” reflecting participants’ experiences of work encroaching of personal time and space. P2 emphasized:

“I work extra hours, and ironically, to protect ‘my KPI’, work-life balance never exists in my dictionary as academia” (P2).

P1, a permanent ECA, added that to avoid being labelled as “not committed” by colleagues, she preferred to use her annual leave to focus on research and publications. She explained:

“My timetable is so packed with classes and meetings that the only way to make progress on research is to take leave. During normal days, I try to keep up with other assigned work and make sure I’m not seen as not committed... I literally apply for leave just to sit down and write or analyse data. This doesn’t happen to me alone; my circle of friends do similar things.” (P1)

Others echoed these sentiments, explaining that raising concerns about workload might result in being assigned even more tasks or being perceived as less dedicated than their peers. Consequently, several ECAs described normalising practices such as staying late, working from home after hours, and sacrificing rest days in order to manage teaching, research, and service simultaneously while avoiding negative judgements about their commitment.

c. Administrative and reporting burden

Participants identified administrative and reporting demands as a major source of strain in their daily work. A substantial amount of time was spent on entering data into online KPI systems, completing forms, uploading evidence, and preparing multiple reports, often across different platforms that were “not merged with one another,” resulting in duplication and delays. The involvement of multiple parties, including department heads and administrative staff, further complicated these processes and increased the time required to complete administrative tasks. Consistent with these experiences, Basheyeva (2025) noted that many educational institutions continue to rely on manual data management systems, which heightens the risk of errors and delays in reporting. Together, these conditions contributed to the perception that administrative duties had become a significant component of workload, frequently at the expense of time for teaching and research. Participants viewed this paperwork burden as an added layer of work tied closely to governance and performance monitoring, yet offering little direct value to teaching quality or research development. P2 echoed:

“Honestly, teaching should play the biggest role in our work. But administrative work is the most tiring and takes up most of my time. My role requires a lot of communication with students and alumni, so it really eats into my time.” (P2)

Similarly, P6 a contract lecturer whom excel in the publications emphasised the administrative load embedded within revenue-focused activities:

“There are so many revenue-focused projects that we get ‘volunteered’ for, and most of them come with heavy administrative work that doesn’t really count towards our own academic development.” (P6)

Together, these accounts illustrate how governance and KPI systems generate significant administrative and reporting requirements that extend beyond routine academic duties, contributing to feelings of overload and limiting the time available for substantive teaching and research work. Interpreted through the JD–R framework, these experiences reflect the growing dominance of job demands within the academic work environment. The accumulation of reporting tasks, publication expectations, and administrative responsibilities created competing pressures on limited time resources, while participants reported relatively few institutional supports to offset these demands. As a result, ECAs experienced an imbalance between increasing performance demands and the resources available to manage them.

Theme 2: Perceived realism and fairness of expectations

Participants frequently questioned the realism and fairness of KPI expectations, particularly in relation to their career stage and available resources. Two sub-themes were identified; (a) realistic versus unrealistic KPI targets; and (b) differences by employment status and gender

a. Realistic versus unrealistic KPI targets

The corporatisation and marketisation of higher education have reshaped academic work, particularly for those in precarious employment, through increased workload, mobility demands, and work–life strain (Izharuddin, 2018). This has made KPIs increasingly unrealistic and challenging to meet. Participants frequently questioned whether KPI targets were appropriate given their experience and the support available. P1, P2, and P6 described the expectations as “targets too high” and “unrealistic for new lecturers,” noting that benchmarks designed for senior lecturers were often applied to junior staff without adjustment. This created a sense that the system did not adequately recognise the developmental needs and constraints of ECAs, reinforcing perceptions of structural unfairness. P2’s experience with her head of department illustrates this tension:

“My KJ (head of department) is a very ‘by the book’ person. She sometimes lacks the human aspect; her only question is, ‘So haven’t you published enough?’ But she doesn’t really help me in any way. It’s demotivating, but it also forces me to be independent.” (P2)

P5, a contract ECA with a professional background in industry who currently holds multiple responsibilities as a programme coordinator and is involved in programme development, emphasised that despite her prior experience, the academic environment remained new and challenging. She argued that institutional expectations did not adequately account for the distinct transition from industry into academia, particularly during the early career stage. Reflecting on the misalignment between performance expectations and lived experience, she stated:

“When I look at the KPI sheet, I honestly wonder who it was designed for, because it doesn’t reflect the reality of being a new lecturer with a full teaching load and limited research funding. On top of that, just so many other uncertain tasks...” (P5)

Similarly, P4 expressed comparable concerns but framed them as an ongoing professional challenge within the academic field. She highlighted the demoralising effect of being assessed against the same performance indicators as senior colleagues who possessed significantly greater research capital and support:

It's demoralising when you're compared directly with senior staff who already have established projects and teams. The KPI numbers might be the same, but the starting point is completely different." (P4)

Overall, this demonstrate that KPI targets were perceived as insufficiently grounded in the realities of academic work. The failure to align expectations with workload conditions and resource constraints reinforced participants' perceptions of unfairness and neglect, particularly when performance measures were implemented without meaningful guidance or developmental input. Rather than serving as motivating benchmarks, KPIs were widely experienced as sources of pressure and emotional strain, shaping a work environment characterised by heightened stress and diminished morale.

b. Differences by employment status and gender

Employment status and gender shaped how KPI pressures were experienced, often creating unequal exposure to risk and support. Contract staff, in particular, perceived a direct link between "KPI tied to contract renewal," which intensified feelings of job insecurity and extra pressure on contract staff. They described situations where KPI sets written into their contracts were "not aligned in the system," creating confusion and a sense of "inequality of risk" compared to permanent colleagues who faced less immediate consequences if targets were not met. Newly working mothers also reported "receiving less support," noting that caregiving responsibilities were rarely considered when workloads and expectations were set, further compounding gendered pressures in meeting performance demands. By contrast, some permanent staff reported greater access to induction and mentoring structures that helped them navigate governance and KPI frameworks. P2, for example, recounted:

"I learnt from PTM. It's an induction course for new lecturer's permanent position. We were introduced with numbers of policies, systems. The course was quite short and not familiar (with the system) because I hadn't experienced it quite yet during that time...but it still helpful." (P2)

P6, who had renewed his contract for a second term, explained how employment status shaped access to support and information:

"Permanent staff get sent to more courses and are more involved in meetings where systems are explained. As contract staff, we are still expected to meet the same KPI or maybe more, but we don't always get the same briefing or mentoring." (P6).

Similarly, P4 observed that permanent staff were afforded greater tolerance for underperformance, whereas contract early career academics faced more severe consequences for unmet KPIs. She described this as an unequal distribution of risk, where a single weak

performance cycle could jeopardise contract renewal. In contrast, P1, a permanent ECA who had recently returned from maternity leave, drew attention to gendered pressures associated with KPI expectations. Despite major changes in her time and energy, performance requirements remained unchanged:

“After coming back from maternity leave, the KPI didn’t change at all. The expectations were exactly the same, but my time and energy were completely different. I don’t see much understanding or adjustment for new mothers.”

Interpreted through the JD–R framework, these accounts suggest that KPI expectations function as intensified job demands that disproportionately affect individuals with fewer institutional resources. Participants highlighted how contract status and caregiving responsibilities limited their capacity to meet publication and performance expectations, thereby increasing feelings of strain and perceived unfairness. These accounts also suggest that gendered caregiving responsibilities may shape how ECAs experience KPI expectations, particularly when institutional performance frameworks do not accommodate periods such as maternity leave or other caregiving responsibilities. Universities may take steps to address these inequalities by offering targeted funding and development opportunities for minority and marginalized groups, including temporary and contract staff. Promotion criteria could also be reassessed to ensure that women and caregivers are not unfairly disadvantaged due to constraints on advancing their publication records or qualifications because of household and family responsibilities (Solomon & Du Plessis, 2023; Mason et al., 2013). Such measures would help mitigate the unequal exposure to risk and support created by current KPI frameworks, where contract staff experience heightened job insecurity and newly working mothers face additional gendered pressures.

Theme 3: Governance structure as enablers or constraints

Analysis of the data revealed three sub-themes, namely: (a) supportive versus restrictive management policies; (b) transparency and communication of expectations; and (c) autonomy in workload negotiation.

a. Supportive versus restrictive management policies

Participants’ accounts highlighted how governance structures, enacted through management practices, could either support or constrain their work under KPI regimes. Many described experiencing predominantly restrictive practices including rigid rules, top-down decisions, and too much bureaucracy that left them feeling disempowered and reluctant to voice concerns. P1 explained:

“It’s better to just keep quiet and not act like you’re an expert... staying silent so you didn’t get much work because there’s no team to rely on and your superior doesn’t really help when issues arise. Some superiors do want to help, many don’t but it really depends.” (P1)

Such reflections suggest that the lack of responsive support and shared decision-making reinforced a culture where ECAs felt they had little influence over how KPIs and workloads were managed, and where raising issues risked further pressure rather than relief.

In contrast, others emphasised the enabling potential of more supportive management, especially where leadership provided concrete assistance rather than merely setting targets. P3, for instance, stressed the importance of broader forms of support:

“I also think it should come in form of resources, assistances not just money and material, but also guidance.” (P3)

Similarly, P5 observed that administrative demands in academia, especially for ECAs, were often excessive. At one end, ECAs encountered bureaucratic and output-driven practices that increased pressure and limited opportunities to voice concerns; at the other, they described the need for approachable leadership, clear guidance, and practical support to help them meet KPI expectations without undermining wellbeing or academic quality.

b. Transparency and communication of expectations

Participants consistently raised concerns about the transparency and communication of performance expectations, particularly around more complex indicators such as KIPs. While they felt that the components for research and teaching were very straightforward, they described KIPs as difficult to understand and not communicated clearly, noting that what counted often “depends on what the superior wants.” P1 explained:

“We have the usual KPIs, and the components for research and teaching are very straightforward. But the KIP is very difficult to understand and often not communicated clearly, because it depends on what the superior wants. That process is not transparent, poorly communicated, and lacks clear planning in terms of how this year’s discussions will inform next year’s KIP. It’s the one that ends up dragging down the other KPIs.” (P1)

Open and transparent decision-making helps lecturers feel informed, valued, and more willing to engage in academic work collaboratively, which ultimately contributes to better educational practice (Ibnu et al., 2024). However, participants’ experiences suggested that such transparency was often lacking. They spoke of unclear performance criteria and the absence of consistent briefings that could have helped ECAs plan their work more effectively and understand institutional expectations. Beyond the technical aspects of KPIs, participants also pointed to broader communication gaps in how important decisions were made and communicated within their institutions. P2, for example, described an incident related to confirmation and promotion:

“There is lack of transparency. Multiple incidents... for example, recently check about confirmation and promotion. In the management meeting, all the candidates should come to the meeting to be interviewed. But it was never widely informed. It should be clear. It should be explained why this situation happened. It should be a normal process to everyone, not just selected one.” (P2)

Experiences like this reinforced perceptions of learning by trial and error, where ECAs discovered critical procedures informally rather than through clear, proactive orientation or briefing sessions. Together, these narratives suggest that limited clarity and inconsistent communication around systems and performance criteria not only hinder effective planning but also contribute to uncertainty and mistrust in governance processes.

c. Autonomy in workload negotiation

Autonomy in workload negotiation emerged as a constrained and uneven aspect of participants' experiences. This pattern echoes existing research, which has established that insecurity is closely linked to uncertain work continuity and limited control over work tasks and responsibilities (Solomon & Du Plessis, 2023). In this study, ECAs generally felt they had limited voice in decisions concerning the distribution of teaching, administrative duties, and additional responsibilities, describing a pattern of routinely "accepting decisions" made at higher levels rather than being actively involved in shaping them. While some participants reported occasional opportunities to "negotiate teaching hours" or request minor adjustments, such instances were described as exceptions rather than standard practice and were often dependent on the goodwill of head of department or dean rather than supported by formal procedures. Many participants distinguished between autonomy within the classroom and autonomy over their broader workload. As P1 explained, decision-making authority was largely confined to teaching practice:

"...50–50 autonomy in handling class, but nothing else. The rest just follow what has been told." (P1)

This illustrates a situation in which ECAs could exercise some professional judgement over how they taught but had little influence over how many classes they were assigned, what additional tasks they were expected to take on, or how their time was allocated across teaching, research, and service responsibilities. Attempts to "push back on tasks" or question these allocations were widely perceived as risky, particularly for those in more vulnerable positions, reinforcing a culture in which compliance with workload decisions was the norm and meaningful negotiation remained limited. P4 further captured this sense of token consultation:

"We are told that workload is 'discussed,' but often it feels more like being informed rather than actually being consulted" (P4)

For contract staff, this lack of bargaining power was experienced even more acutely. P5 noted that although staff were formally told they had the right to refuse tasks, in practice such refusals carried perceived risks, particularly in relation to contract renewal:

"I can suggest small changes like swapping a class time, but when it comes to the overall number of courses or committees, it's already decided. We just follow. They told that we have the rights to refuse. But, what do you think for contract like us? It's going to give bad impact when renewing the contract." (P5)

Taken together, these accounts illustrate how governance structures can operate either as enabling resources or as additional constraints within the academic work environment. While some participants acknowledged the potential value of supportive leadership, clear communication, and practical guidance, many described governance practices that limited their voice in decision-making and constrained their ability to negotiate workloads. Interpreted through the JD–R framework, such governance arrangements may function either as organisational resources that support ECAs in navigating institutional expectations or as additional job demands when policies are experienced as rigid, opaque, or excessively bureaucratic. When supportive leadership and transparent communication were absent, participants reported greater uncertainty and reduced autonomy, reinforcing the imbalance between institutional performance demands and the resources available to manage them.

5. Conclusion, Implication and Recommendation

Framing the findings through JD-R theory clarifies why governance and KPI systems were experienced by ECAs as both overwhelming and, in some cases, potentially enabling. JD–R theory explains work experiences as shaped by the balance between job demands such as sustained effort and emotional strain and job resources, which support goal attainment and development (Bakker & Demerouti, 2007). In this study, KPI-driven task prioritisation, extended working hours, heavy administrative and reporting requirements, unclear performance expectations, and non-negotiable workloads functioned as major job demands, particularly when targets were perceived as unrealistic or when employment status and gender increased the risks associated with underperformance. At the same time, job resources were unevenly distributed: supportive leadership, clear communication, induction programmes, mentoring, and flexible workload arrangements helped some ECAs manage these pressures more effectively, while contract staff and newly working mothers often reported limited access to such support. In particular, the experiences of newly returned mothers highlighted in this study suggest that gendered caregiving responsibilities constitute an important but underexplored dimension of KPI-driven academic performance cultures, meriting focused, gender-sensitive research designs in future studies. From a JD–R perspective, caregiving responsibilities may further reduce the resources available to ECAs for meeting institutional performance demands, thereby intensifying the imbalance between job demands and available resources. Consistent with JD–R theory, this imbalance between high demands and insufficient resources contributed to stress, work–life conflict, and disengagement; however, the findings also suggest that strengthening institutional resources could shift KPI governance from being primarily draining to more developmental in nature.

Based on these findings, several implications and recommendations arise. Institutions should review KPI frameworks for career stage sensitivity, ensuring differentiated expectations, scaffolded targets, and explicit recognition of transition periods, caregiving, and disciplinary variation. Governance processes need greater transparency through systematic, accessible briefings on policies, promotion, and appraisal, reducing reliance on informal “trial and error” learning. Workload models should incorporate mechanisms for genuine negotiation, especially for ECAs and contract staff, and reduce redundant administrative reporting by streamlining systems. Finally, leadership development and mentoring programmes should emphasise

supportive, dialogic management that combines accountability with resources, guidance, and advocacy, so that KPIs become tools for capacity building rather than drivers of chronic strain.

Future research should deepen and extend these insights in several ways. Comparative studies across different institutional types, disciplines, and national systems could clarify how specific governance configurations and funding regimes shape the demand–resource balance for ECAs, including cross-country analyses informed by JD-R theory. Longitudinal qualitative or mixed-methods designs could track how ECAs’ experiences of governance, KPIs, and wellbeing evolve over time, particularly across key transition points such as confirmation, promotion, or shifts from contract to permanent roles. Further work is also needed on intersectional experiences examining how gender, caregiving status, contract type, ethnicity, and field of study interact to produce distinct configurations of risk and resilience in performance cultures. In particular, the experiences of newly returned mothers highlighted in this study suggest that gendered caregiving responsibilities may represent an important but underexplored dimension of academic performance cultures. Finally, intervention-focused research that pilots and evaluates concrete changes (e.g., revised KPI frameworks, mentoring schemes, workload negotiation mechanisms) would offer much-needed evidence on what kinds of institutional reforms most effectively support sustainable and equitable ECAs’ trajectories.

AI Declarations

AI-assisted tools (ChatGPT, Gemini, and Quillbot) were used during manuscript preparation for language editing and grammar checking. These tools were not used for data analysis, interpretation of findings, or generation of research conclusions. All analytical decisions and interpretations were conducted by the authors.

References

- Abreu, M., & Grinevich, V. (2024). The entrepreneurial university: Strategies, processes, and competing goals. *The Journal of Technology Transfer*, 49, 1991–2034. <https://doi.org/10.1007/s10961-024-10085-7>
- Ani Marlia, M., Fahmy, R., Lukito, H., & Games, D. (2025). Explicating Collegiality and Change Management in Neoliberalism during The Dynamics of Higher Education Institutions: A Systematic Literature Review using PRISMA Checklist. *F1000Research*, 13, 237.
- Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of managerial psychology*, 22(3), 309-328.
- Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The JD–R approach. *Annual review of organizational psychology and organizational behavior*, 1(2014), 389-411
- Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands–resources theory: Ten years later. *Annual review of organizational psychology and organizational behavior*, 10(1), 25-53.
- Basheyeva, Z., Oralkhanov, D., Orazov, T., & Soltan, G. (2025). Development and implementation of an automated web-based kpi management and dashboard system at astana it university. *Scientific Journal of Astana IT University*, 46-62. <https://doi.org/10.37943/23mzin6824>
- British Educational Research Association. (2018). Ethical guidelines for educational research (4th ed.). *British Educational Research Association*. <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018>
- Bone, K. D. (2021). Cruel optimism and precarious employment: The crisis ordinariness of academic work. *Journal of Business Ethics*, 174(2), 275–290. <https://doi.org/10.1007/s10551-020-04605-2>
- Bozzon, R., Murgia, A., Poggio, B., and Rapetti, E. (2017). Work–life interferences in the early stages of academic careers: The case of precarious researchers in Italy. *European Educational Research Journal* 16, 332–351. <https://doi.org/10.1177/1474904116669364>
- Burton, A., & Bowman, S. (2022). The rise of the academic precariat: An analysis of contingent faculty labour in UK higher education. *Work, Employment and Society*, 36(3), 517–533. <https://doi.org/10.1177/0950017021102641>
- Deem, R., Hillyard, S., & Reed, M. (2007). *Knowledge, higher education, and the new managerialism: The contemporary university and the maintenance of inequality*. Oxford University Press.
- Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The Job Demands-Resources Model of Burnout. *Journal of Applied psychology*, 86(3), 499.
- Franco-Santos, M., & Doherty, N. (2017). Performance management and well-being: A close look at the changing nature of the UK higher education workplace. *The International Journal*

of *Human Resource Management*, 28(16), 2319–2350. <https://doi.org/10.1080/09585192.2015.1130638>

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>

Harford, J. (2021). The Irish state and higher education policy: From élitism to massification. *Irish Educational Studies*, 40(3), 329–347. <https://doi.org/10.1080/03323315.2021.1977983>

Henkel, M. (2004). Current Science Policies and Their Implications for the Formation and Maintenance of Academic Identity. *Higher Education Policy*, 17(2): 167–182. doi: 10.1057/palgrave.hep.8300049.

Hollywood, A., McCarthy, D., Murphy, E., & Shevlin, M. (2020). Illuminating the path: The career journey of early career academics in STEM disciplines in higher education. *Studies in Higher Education*, 45(11), 2311–2325. <https://doi.org/10.1080/03075079.2019.1634660>

Huang, Q. (2024). From Ivory Towers to Corporate Ladders: The Impact of Managerialism on Academic Profession. *Adult and Higher Education*, 6(1), 159-165. <https://doi.org/10.23999/aduhe.2024.060124>

Ibnu, M., Ritonga, K. S., Hanafi, M. H., Maulana, A., & Darmansah, T. (2024). Peran Manajemen Perkantoran Dalam Meningkatkan Kinerja Tenaga Pendidik di MTS Muallimin Univa Medan. *Jurnal Budi Pekerti Agama Islam*, 2(3), 288-297. <https://doi.org/10.61132/jbpai.v2i3.454>

Izharuddin, A. (2018). Precarious intellectuals: The freelance academic in Malaysian higher education. *Kajian Malaysia* 36, 1–20. <http://doi.org/10.21315/km2018.36.2.1>

Javed, Y. and Alenezi, M. (2023). A Case Study on Sustainable Quality Assurance in Higher Education. *Sustainability*, 15(10), 8136.

Laudel, G., & J. Gläser. 2008. From Apprentice to Colleague: The Metamorphosis of Early Career Researchers. *Higher Education*, 55 (3): 387–406. <https://doi.org/10.1007/s10734-007-9063-7>.

Lee, A., Dennis, C., & Campbell, P. (2021). Nature’s guide for mentors. *Nature*, 447(7146), 791–797. <https://doi.org/10.1038/447791a>

Li, H., Wang, W., & Chen, J. (2025). The double-edged sword: Performance-based pay and its impact on work-life balance and burnout among university faculty. *Higher Education Research & Development*, 44(1), 1–18. <https://doi.org/10.1080/07294360.2025>

Lynch, K., & Ivancheva, M. (2015). Academic freedom and the commercialisation of the university: A critical analysis. *Higher Education*, 70(2), 289–309. <https://doi.org/10.1007/s10734-014-9842-m>

Mason, M. A., Wolfinger, N. H., and Goulden, M. (2013). *Do babies matter? Gender and family in the ivory tower*. Rutgers University Press.

McKenzie, J. (2018). Navigating neoliberalism: The challenge for early career academics. *Higher Education Research & Development*, 37(6), 1109–1123. <https://doi.org/10.1080/07294360.2018.1487841>

Naidoo-Chetty, M., & Du Plessis, M. (2021). Job demands and Job Resources of Academics in Higher Education. *Frontiers in psychology*, 12, 631171.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. U.S. Department of Health, Education, and Welfare.

Power, M. (1997). *The audit society: Rituals of verification*. Oxford University Press.

Slaughter, S., & Rhoades, G. (2012). *Academic capitalism and the new economy: Markets, state, and higher education* (2nd ed.). The Johns Hopkins University Press.

Solomon, S., & Du Plessis, M. (2023, May). Experiences of precarious work within higher education institutions: A qualitative evidence synthesis. In *Frontiers in Education* (Vol. 8, p. 960649). Frontiers Media SA.

Spina, N., Gill, J., & Tonna, M. (2020). The precarity of academic work in higher education: The case of casual academics in Australia. *Higher Education Research & Development*, 39(5), 1000–1013. <https://doi.org/10.1080/07294360.2020.1712681>

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics and the academy*. Routledge.

Teichler, U. (2019). *Steering in a modern higher education system: The need for better balances between conflicting needs and expectations* (INCHER Working Paper No. 11). International Centre for Higher Education Research.

Wan, A., Azman, Z., & Sirat, M. (2022). Institutional productivity and the influence of neoliberal and new public management on Malaysian public universities. *Malaysian Journal of Higher Education*, 8(1), 1–18.

Zhuang, T., & Lin, M. (2025). Early-career academics' grappling with multi-dimensional duties in the one-dimensional academic world: a temporal sociological lens. *Cambridge Journal of Education*, 1-19

Appendix A: Interview Protocol

Context	Sample of the Interview Questions
Opening (Informed consent and participation)	Thank you for taking the time to participate in this interview. Before we begin, I would like to briefly explain the purpose of this session and your rights as a participant. This interview is part of a research project that aims to understand how governance systems and KPI expectations shape ECA's daily academic work, workload, and priorities. • Everything you share today will be treated with strict privacy and confidentiality. Your name and any identifying information will not appear in any reports, publications, or presentations. All data will be anonymised, and only the researcher will have access to the raw interview files. • Your participation in this study is completely voluntary. You have the right to skip, pause the interview, or withdraw from the study at any time. If you choose to withdraw, your data will not be used. • With your permission, I would like to audio-record this interview. The recording will help ensure accuracy during transcription and analysis. The audio files will be securely stored and will not be shared with anyone outside the research team. Before we proceed, may I confirm the following: 1. Do you understand the purpose of this study? 2. Do you agree to participate voluntarily? 3. Do you consent to the audio recording of this interview? If you agree, we can begin the recording and start the interview.
Background and context	1. Could you briefly describe your current role and responsibilities? 2. How would you describe a typical workday or workweek?
Perception of governance	1. How do you understand the governance structure in your institution or faculty? 2. How would you describe your experience interacting with institutional policies, rules, or decision-making processes? 3. In your opinion, how does governance affect your work as an ECA? ○ Can you elaborate further?

Influence of KPIs and metrics	1. What KPIs or performance metrics are you expected to meet? - o Can you give any example? 2. How do these metrics shape your daily tasks or decisions? 3. How realistic or achievable do you find these KPIs?
Emotional & motivational experiences	1. How you balance teaching, research, and administrative work? 2. Do you feel supported by your institution when working to meet these expectations?
Closing questions	1. Is there anything else you would like to add that we haven't discussed?

Appendix B: Emergence of New Codes and Thematic Saturation by Interview

Interview	New Codes Identified	Cumulative Codes
P1	18	18
P2	12	30
P3	8	38
P4	5	43
P5	3	46
P6	0	46

Refining Qualitative Research Instruments Through Participant Involvement: An Interview Protocol Refinement Process for ELT Curriculum Research in Uzbekistan

ABSTRACT

Authors: Nozimakhon Uralova^{1,2} and Sharifah Intan Sharina Syed Abdullah^{1*}

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

²English Language Teaching and Educational Management, Central Asian University, 111221, Milliy Bog St, Tashkent, Uzbekistan.

*Corresponding author:

sharifahintansharina@upm.edu.my

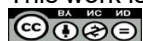
Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This paper describes an Interview Protocol Refinement (IPR) process conducted as a preparatory phase for a qualitative research that explores the perceptions of Uzbekistan's university lecturers regarding curriculum development and evaluation methods of the English Language Teaching (ELT) curriculum implemented at private universities. The purpose of this qualitative phase is to systematically refine semi-structured interview protocol to maximize the trustworthiness and authenticity of the data collection in the main study. This article reports and describes on four stages of the IPR process: (1) brainstorming the questions and aligning them with the research questions using the matrix, (2) conducting external review with the experts, (3) conducting internal review with university lecturers, (4) piloting interview questions study as a final stage of the IPR process. This paper demonstrates how the involvement of the perspectives of research participants and experts enhanced the rigor and quality of the research instrument. The paper provides a detailed examination of each stage, specific weaknesses in question design, alignment of the questions with objectives of the main study, appropriate sequencing, word choice which are systematically addressed through revision stages. The preparatory phase clearly illustrates that qualitative research is a cyclical process and that stakeholder engagement in an instrument development stage could improve the quality of subsequent data collection. The findings might serve as contribution to further methodological discussions and systematic approaches for qualitative instrument development, offering a practical guidance for research designing semi-structured interview protocols.

Keywords: Pilot Studies, Interview Protocol, Review, Research Instrument, Researcher Bias, University, Lecturers, ELT, Curriculum Evaluation, Uzbekistan

1. Introduction

1.1. Changes in Higher Education Sector in Uzbekistan

Uzbekistan is the country located in Central Asia with a population of approximately 38 million people. The country was a part of the Soviet Union till 1991. After gaining its independence, investments in education has become Uzbekistan's top priority in enhancing its economic and social development (World Bank Report, 2023). However, at the beginning, higher education experienced multiple challenges such as insufficient funding, issues in educational quality, out-dated programs and curriculum that failed to align with the needs of evolving economic and industrial demands of the country. This subsequently resulted in low education enrolment and interest in local universities among Uzbek youth (World Bank Report, 2023). Nevertheless, population growth and changing economic conditions of the country, steadily increased the demand for educational reforms in Higher Education Institutions (Ruziyev & Burkhanov, 2018). During the Soviet Union era, there were only 43 Higher Educational Institutions (HEIs), but since independence this number significantly grew to 222, including private and branches of international universities (UzDaily, 2025). A Presidential resolution endorsed in 2017, PQ-2909, outlined key provisions to develop the higher education system in Uzbekistan. The resolution details the measures for modernizing infrastructure, reforming curriculum, enhancing faculty qualifications, and establishing cooperation with leading international universities (Lex.uz, 2017).

The core changes in the higher education sector began with the establishment of a systematic reform "The Concept of Development of Higher Education system of the Republic of Uzbekistan until 2030", the Decree of the President of the Republic of Uzbekistan No. PP-4391, dated 2019 (Lex.uz, 2019). This strategic document has brought many positive changes in the higher educational system (Rasulov, 2024). The universities have started transitioning to a credit-modular system based on the European Credit Transfer and accumulation System (ECTS), 40 local universities have been granted academic and financial autonomy which allowed them to make their decisions independently regarding their programs, curriculum, financial and qualifications requirements (Rasulov, 2024). Another significant reform was the change in licensing process for private educational establishments that has led to a boom of private universities. The pioneer private university was Akfa University (now Central Asian University), founded in 2019. Currently, the number of private universities has reached 70, demonstrating a rapid expansion of private higher education institutions in the country (Erasmus+ National Office in Uzbekistan, 2024).

On the other side of the spectrum, this surge necessitates maintaining rigorous quality standards in private universities during this short period of expansion. All private universities face a critical need to develop effective programs and curricula due to heightened competition and increasing interest in the private education sector. As private universities are newly established organizations, the majority of them have not enhanced their robust quality control and quality assurance mechanisms in terms of curriculum development and evaluation yet (Shadiev et al., 2025). Moreover, most university programs do not possess their own curriculum which is tailored to the context of a private university. Curriculum is mainly adopted or adapted from public universities which are not quite aligned with the requirements in the private educational sector and do not meet the demands of the stakeholders who choose private universities.

Another significant issue identified is that universities in Uzbekistan implement "top-down" to "bottom-up" curriculum development and evaluation approach (Bezborodova, 2022), which fails to provide "clear guidance", authentic communication and support to the main stakeholders (Bezborodova & Radjabzade, 2022; Zhang et al. (2025) who are university lecturers and students. The curriculum in HEIs is designed by "the authority" and "limited pool of persons" (Mandel, 2015). Lack of communication between "top" - who are curriculum developers, experts or Ministry of Higher Education and "bottom" who are lecturers, the main implementers, "actors" (Zhao, 2011) in curriculum

changes leads to the gaps in curriculum development and implementation (Khan et al., 2023). As a result, another challenge has been faced by private universities - maintaining the balance of the needs among various stakeholders (students, lecturers, parents, curriculum developers, Ministry of Higher Education, employers) in designing and evaluating the curriculum. Each group of stakeholders might prioritize different outcomes which makes it challenging to establish a consistent and clear process of curriculum development and evaluation (Khan et al., 2023). Implementing only “top-down” curriculum development approaches have brought misalignments between content sequencing, assessment strategies, and curriculum outcomes and objectives (Cheema, 2021; Hafeez et al., 2021).

Most studies highlight the importance of lecturers’ role in curriculum development and evaluation (Louvel, 2013; McGrath et al., 2016; Micallef, 2025; Stephens et al., 2008; Van der Heijden et al., 2015). Lecturers are defined as “agents of change” (Nieveen, 2011; Priestley, 2011), who can contribute significantly to the continuous improvement of curriculum alongside curriculum experts. This paper describes the first phase of qualitative research on exploring the perspectives of university lecturers regarding English Language Teaching (ELT) pre-service teacher education curriculum development and evaluation at private universities in Uzbekistan. The first phase of the study was focused on Interview Protocol Refinement and pilot study that prioritizes the perspectives of “bottom-up” curriculum implementers (lecturers) who directly work with and are aware of the needs of the target population of the ELT curriculum (pre-service teachers). This phase of the research actively involves lecturers into the research instrument development and improvement, revealing the expectations of stakeholders and qualities of a market-responsive curriculum that meets the needs of pre-service teachers and employers.

1.2. The focus of the Primary Study

This qualitative research explores the perceptions of lecturers who are described as the “product implementers” (Zhao, 2011), and those who can share “information-rich” (Patton, 2015) data and meaningful perspectives because of being constantly engaged with the curriculum. The primary research explores two important aspects of the ELT curriculum for pre-service teacher education. First, the quality and essential attributes of the well-designed ELT curriculum. Second, how the ELT curriculum should be evaluated in a context of Higher Educational Institutions, focusing on the criteria and approaches of the curriculum evaluation process.

The initial phase of the research serves to improve the core research instrument of the study - the interview protocol, and consists of the following aims:

- to enhance the interview protocol of the research by involving university lecturers into the internal review and pilot study stages
- to apply the changes and modify the interview questions after each phase

2. Literature Review

2.1. Integrating Lecturer Insights into Curricular Governance

In the context of HEIs, curriculum development is defined as a comprehensive decision-making process and the “compilation of diverse individual contributions” that needs constant improvement and innovation (Gordon et al., 2019). Karkkainen (2012) emphasizes that lecturers should be involved to review curriculum and become curriculum innovators that further enhance the quality of teaching and learning. He also strongly suggests that change in curriculum could be a turning point for innovation in education. Ball (1990) describes lecturers as “knowledgeable people” with extensive understanding of the field due to their professional role, expertise and experience. Many studies found the benefits of collecting data from lecturers to enhance the curriculum in HEIs, leading to effective achievement of educational reforms (Carl, 2009, Fullan, 1991; Handler, 2010; Van der Heijden et al., 2015).

Datnow & Hubbard (2016) highlight that collecting reliable data from lecturers enables educational leaders to obtain authentic and comprehensive information regarding the program and the main stakeholders such as students, parents and industry representatives. This, in turn, leads to identifying strengths and weaknesses, and areas for improvement of curriculum more accurately. As a result, educational leaders can gain practical evidence, develop targeted strategies to address the needs of students, and improve the quality of the program (Halverson et al., 2007). McGrath et al., (2016) point out that HEIs leadership who often ignore the insights of lecturers has a high possibility of delivering curriculum that fails to meet contemporary industry expectations and student needs. Therefore, curriculum development, revision and evaluation cycles should incorporate formal mechanisms for lecturers' input and contributions (Louvel, 2013; Mandel, 2015).

2.2. Curriculum evaluation in English Language Teaching Pre-service Teacher Education

According to Peacock (2016), it is crucial for every teacher-training program to have a system for regular internal evaluation. Rea-Dickins and Germaine (1998) also emphasize “the importance of having systematic evaluation at the heart of a programme”. Peacock (2016) mentions that there seem to be only 'few detailed descriptions' on how to evaluate current ELT programs. To determine if a curriculum is equipped with all necessary skills and to identify the shortcomings of ELT programs, there is a need for a model that can be adapted to the context and can evaluate the effectiveness of teacher education programs. Having an effective ELT program evaluation can contribute to “program improvement” (Lynch, 2003). Yang (2009) suggests that there should be a clearly stated philosophy and that “program content should reflect it”.

2.3. Benefits of IPR and Piloting Studies to Enhance the Research Instrument

The quality of obtained data depends significantly on well-designed research instruments. Every novice and emerging qualitative researcher needs support and guidance in designing effective research instruments that help “produce valid, reliable, unbiased, and complete” results (Cohen, 2003) to ensure methodological rigor. When qualitative research employs in-depth interviews as the primary research instrument, researchers should make sure that interview protocols are “carefully designed and structured” “to enhance validity and reliability” and to reduce bias (Castillo-Montoya, 2016; Merriam & Tisdell, 2016). A comprehensive and well-designed interview protocol leads to obtaining insightful and information rich data while equipping a researcher with clear structure, consistency and guidance throughout the data collection process (Jones et al., 2013; Miles et al., 2014).

Qualitative studies focusing on exploring perceptions of individuals are often complex and challenging to investigate through interviews due to a variety of factors such as not thinking deeply or explicitly about their beliefs (Braaten, 2020). Therefore, researchers should plan their interviewing approaches purposefully to provide interviewees with adequate space and prompting that gives opportunity for deeper reflection. Additionally, Braaten et al., (2020) suggests researchers to systematically gather data by “organizing the interview in a way that will allow them to answer their research questions”.

Castillo-Montoya (2016) highlights the benefits of creating a matrix prior to designing an interview protocol that demonstrates the alignment of interview questions and research questions. This alignment “can increase the utility of interview questions, confirm their purpose” and visually demonstrate the value of each question, keeping necessary questions and eliminating unnecessary ones (Castillo-Montoya, 2016). This process is considered as an initial step of Interview Protocol Refinement Framework (IPR). Thus, to enhance reliability of interview protocols, implementing Interview Protocol Refinement Framework suggested by Castillo-Montoya (2016) could serve as a tool of “rigorous steps to interview protocols and ensure their congruency” (Jones et al., 2014). The IPR framework includes four essential stages (Castillo-Montoya, 2016):

- Stage 1: Aligning interview questions with research questions
- Stage 2: Constructing an inquiry-based conversation
- Stage 3: Receiving feedback on interview protocols
- Stage 4: Piloting the interview protocol

The piloting stage of the IPR strengthens the quality of the questions by enhancing their clarity, and identifying possible issues with timing. Kvale (2007) claims that pilot studies can assist to detect weaknesses and limitations in interview design, providing an opportunity for a researcher to make adjustments prior to applying in the main study (Kim, 2010). Like the other three stages of the IPR, the piloting stage serves to refine research instruments through testing (Bickman & Rog, 2009; Kim, 2010; Yin, 2014) and to evaluate the feasibility of recruitment protocols (Bickman & Rog, 2009; Secomb & Smith, 2011; Van Wijk & Harrison, 2013). In other words, this final stage of refinement allows researchers to test whether participant selection criteria are appropriate and whether effective participants who can help answer the research questions are being reached (Maxwell, 2013). This stage also reveals if the order of the questions is effective and appropriately placed or needs modification (Merriam, 2009), and provides information about the complexity of the questions and whether participants are able to answer them (Maxwell, 2013).

Moreover, this stage generates preliminary data (Janghorban et al., 2014) and increases confidence of a novice researcher in conducting in-depth interviews (Ismail et al., 2018; Janghorban et al., 2014). Stage 4 of the IPR identifies potential improvements and makes final revisions of interview protocols (Maxwell, 2013). The IPR framework is defined as an effective methodology for generating robust interview protocols that serves to collect “rich, focused, and meaningful data that captures participants’ lived experiences” (Castillo-Montoya, 2016) that fully aligns with the purpose of this research. This qualitative research focuses on lecturer perceptions regarding ELT curriculum development and evaluation, emphasizing their experiences and key contributions toward enhancing the curriculum at private universities. The lecturers were also actively involved in the IPR process, which constitutes the initial phase of the research.

2. Methodology

2.1. The Stages of the Pilot Study

To enhance the quality of the research instrument, qualitative research employed the IPR framework as the preliminary phase of the data collection (Figure 1).

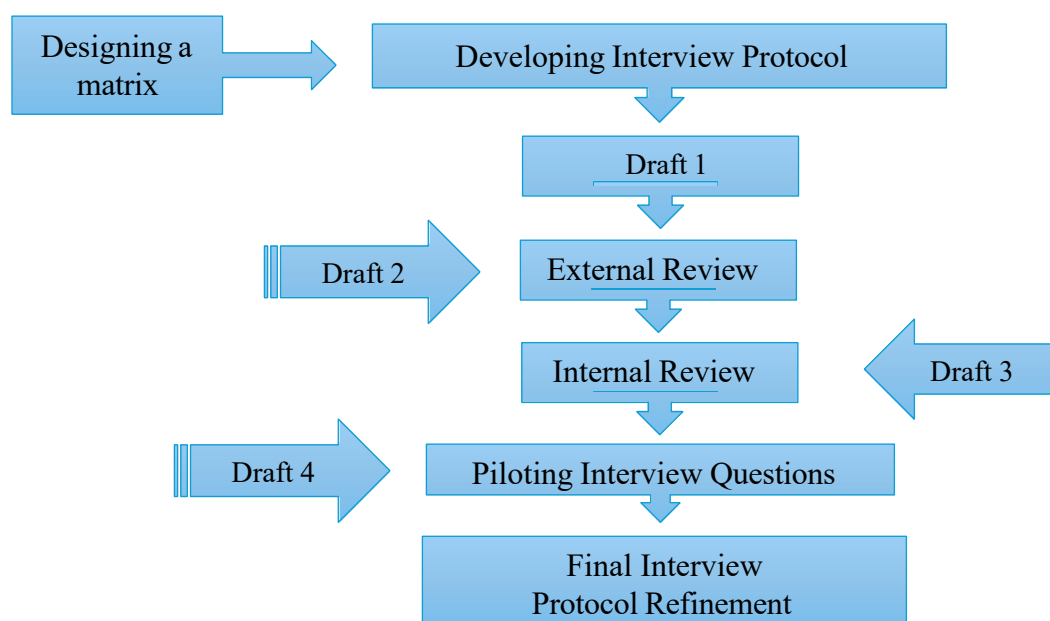


Figure 1. The stages of the IPR

This qualitative research explores the perceptions and perspectives of the main "actors" (Zhao, 2011) in ELT curriculum development and evaluation. Research question 1 focuses on lecturers' perspectives on the ELT curriculum content and components, while research questions 2 and 3 aim to explore lecturers' insights on curriculum evaluation quality and approaches they consider to be effective. In-depth semi-structured interviews serve as the core data collection method. The study modifies an existing framework by integrating a robust review Leadership Competencies process that involves lecturers as reviewers as well as pilot study participants who help enhance the research instrument. Prior to constructing the interview protocol, I addressed the following questions (Bryman, 2012; Creswell, 2018; Kvale, 2007) that helped in preparing the qualitative interview protocol:

- What do I need to know to answer my research questions?
- What are possible topics that need to be covered in the interview questions?
- What do I know about the context?
- Who are the participants?
- What do I know about the setting?
- What type of question could be useful to answer the research questions: general or specific?
- What language should be used?
- How can I avoid the questions that are leading and biased?

Having answered these questions, I compiled a list of topics related to ELT curriculum development and evaluation. The topics were logically organized, progressing from general to specific - beginning with ELT methodology and assessment, then moving to curriculum design, finally addressing curriculum evaluation. Subsequently, I constructed the questions that "flow reasonably well" with the topics, acknowledging the possibility that "the order of the questions could be changed during the actual interview" (Bryman, 2012). This process was a preliminary stage - brainstorming of the questions that could help me answer the research questions. I ensured that the language used in the questions was "comprehensible and relevant to the people I am interviewing" (Bryman, 2012; Kvale, 2007), who are lecturers of the ELT program and familiar with ELT methodology, language testing and assessment, curriculum design and evaluation. Following this, I created a matrix that "mapped interview questions onto research questions" (Castillo-Montoya, 2016), visually demonstrating the alignment of each interview question with the research questions. During this stage, the questions were filtered, eliminating those which are unlikely to contribute to collecting meaningful data (Castillo-Montoya, 2016). The original matrix suggested by Castillo-Montoya (2016) included interview questions, background information and research questions (Table 1).

Table 1. Sample Matrix (Castillo-Montoya, 2016)

	Background Information	Research Question 1	Research Question 2	Research Question 3
Interview Q1	X			
Interview Q2	X			
Interview Q3		X		
Interview Q4		X	X	X

However, the matrix I constructed contains additional elements. It includes three research questions, topics addressed in those research questions, interview questions and rationales (Table 2). Integrating rationales explicitly indicates my position as a researcher, demonstrating one of the qualities of "a good qualitative inquirer" (Hammersley & Atkinson, 1995). This matrix helps maintain the trustworthiness

of the research procedure, increasing credibility and dependability (Lincoln & Guba, 1985). It serves to document every decision made and encourages researcher reflexivity (Maxwell, 2013). The matrix also assists to track the balance of the questions related to each research question, identify gaps, evaluate, adjust, add or eliminate certain questions. The interview protocol was divided into three sections: opening questions, core questions, and closing questions. The opening section included five questions that were used for rapport-building and essential information collection that covers facesheet information, and background information. The core questions section consists of three sub-sections, each reflecting one of the three research question:

- i. Curriculum experience
- ii. Perceptions of effective curriculum
- iii. Criteria and approaches for curriculum evaluation

Each sub-section contains 2 to 4 main questions and 2 to 5 prompts and follow-up questions. The final section includes three questions on challenges, identification of barriers, and a question that encourages sharing additional reflection and insights of participants.

Table 2. Sample Matrix

Topics covered	RQs	Core Questions	Prompts & follow-up questions	Rationale
ELT curriculum, content of the curriculum, assessment practices	RQ 1	What key elements do you believe should be part of an effective ELT curriculum? (to meet the needs of pre-service teachers)	<i>a. Could you describe a particular component of the ELT curriculum that is the most effective and/or ineffective?</i> <i>b. To make the curriculum more effective, what courses do you consider the most useful for preparing future English language teachers?</i> <i>c. What specific assessment methods (e.g., portfolio, observed lessons, self-reflection journals) do you believe are most effective for evaluating a student teacher's readiness?</i>	The questions progress from general to specific, highlighting what elements should be inside the curriculum. The first follow-up question encourages obtaining broad and diverse perspectives of participants. Next questions are more specific, asking about the necessary courses, assessment, and skills that should be integrated inside the ELT curriculum that prepares pre-service teachers for real teaching context

The next stage of the IPR was significant in terms of detecting bias, leading questions and assessing the effectiveness of the interview protocol externally and internally. This review process involved six experts,

three external and three internal reviewers. Selected external reviewers were the researchers with more than 10 years of experience and expertise in the field of educational research and ELT curriculum development. The experts hold PhD degrees in education and are university lecturers with “a track record of research and published work in the relevant field” (Scorsolini-Comin, 2020). The experts were selected according to the following criteria:

Academic qualifications: Holding PhD Degree in Education

- Extensive experience in research: More than 15 years of experience in educational research
- Subject matter experience: Having knowledge and expertise in the field of Education, ELT or Curriculum Development and Evaluation
- Employment: Currently working as a lecturer in ELT or Language Education Department
- Institutional diversity: Being from different universities to avoid bias and maintain objectivity and neutrality

Table 3 provides brief information about the external experts.

Table 3. External Experts

External Experts	Degree/Position	Experience	University	Field of expertise
Expert 1	PhD, Associate Professor and Deputy Head of Research	15	New Uzbekistan University, Uzbekistan	ELT, Curriculum Development
Expert 2	PhD, Lecturer	11	The Hong Kong Polytechnic University, China	ELT
Expert 3	PhD, Lecturer	15	McGill University, Canada	Education

The experts suggested the characteristics of a good qualitative interview and designed the criteria for evaluation. The core purpose of the review was to examine the overall quality of the questions. They reviewed the protocol based on the following criteria (Bryman, 2012; Castillo-Montoya, 2016; Coker & Akande, 2025; Maxwell, 2013; Kvale, 2007):

1. Clarity of wording
2. Relevance of questions
3. Coverage of questions
4. Sequencing and flow
5. Language appropriateness
6. Ethical considerations
7. Overall suitability of the protocol

The protocol was also reviewed in terms of methodological alignment to ensure the questions are consistent with the selected qualitative approach, allowed for detailed and rich responses, encouraged storytelling and examples and can be easily adapted or reordered during the interview if it is appropriate.

Internal reviewers were selected based on the following criteria:

- Academic Qualifications: MA in Teaching English to Speakers of Other Languages (TESOL), PhD in TESOL or Curriculum and Instruction
- Subject matter expertise: Knowledge in the field of ELT and Curriculum Development
- Methodological experience: Experience in qualitative research

- Employment: Currently working as a lecturer in an ELT department

The internal reviewers involved in the process were three lecturers who work at a private university ELT program, each with more than five years of teaching experience in HEIs. All three reviewers were lecturers who work at the ELT department. In addition to their teaching roles, two of the reviewers had experience in ELT curriculum development and are currently pursuing their PhDs in the field of Curriculum and Instruction and TESOL. The third reviewer holds a PhD in TESOL and is currently working as a Senior Lecturer. Table 5 provides detailed information about the internal reviewers.

Table 5. Internal Reviewers and Pilot Study Participants

Internal Reviewers	Qualifications/Position	Experience in research	University	Field of expertise
Reviewer 1	PhD in TESOL, Senior Lecturer	15	Kimyo International University, Uzbekistan	ELT, TESOL
Reviewer 2	PhD candidate in Curriculum and Instruction, Senior Lecturer	7	New Uzbekistan University, Uzbekistan	ELT, Curriculum and Instruction
Reviewer 3	PhD candidate in TESOL, Senior Lecturer	8	Webster University in Tashkent	ELT, TESOL, Curriculum and Instruction

Reviewing and getting feedback from fellow colleagues led to modifications and improvements in the quality of questions. I constructed the checklist (Table 4) based on the recommendations by Bryman, 2012; Maxwell, 2013; and Patton, 2015. The checklist consisted of seven criteria, and questions and prompts were asked to be examined using the checklist and the reviewers' expertise as lecturers. The main purpose of the review was to examine the questions in detail and determine suitability of providing rich information. Using the checklist, reviewers provided comments on whether the questions are not too broad or vague, follow logical sequence, language appropriateness, neutrality of questions, practical considerations, and ethical considerations.

Table 4. Sample checklist (Bryman, 2012; Maxwell, 2013; and Patton, 2015)

Research Alignment	• Questions directly contribute to answering the research questions • Questions cover all key aspects needed to address research objectives • Questions have clear purpose (descriptive, exploratory, explanatory) • Questions are sufficiently open but not too broad or vague • Questions follow a logical sequence or progression
Language Appropriateness	• Language is comprehensible to target participants (avoids jargon or technical terms unless appropriate)

	• Vocabulary is relevant to participants' experiences and context • Questions are culturally appropriate and sensitive • Language is clear and unambiguous • Questions are concise and focused (not overly complex or double-barreled) • Reading level is appropriate for participants
Neutrality and Balance	• Questions avoid leading participants toward specific answers • Questions don't contain implicit assumptions • Questions are free from value-laden terminology • Questions allow for a range of possible responses (positive, negative, neutral) • Questions don't suggest "correct" or socially desirable answers
Ethical Considerations	• Questions respect participant privacy and dignity • Potentially sensitive topics are approached carefully • Questions don't create undue psychological stress • Participants can reasonably decline to answer without discomfort

The final stage of the IPR was piloting the interview protocol, once more involving lecturers in the study. The primary objective of the pilot study was to complete refinement of the interview protocol. Participants for this pilot stage were recruited through purposive sampling. The recruitment process proceeded through the announcement of the pilot study on social media (LinkedIn page) professional network channel. 15 university lecturers expressed their initial interest and requested full information about the study. The detailed information sheets about the study and a consent form were sent to 10 out of 15 who were lecturers currently working at a private confirmed their participation. Finally, four of them were selected based on the following inclusion criteria. university. After receiving the full information about the study and the consent form, 6 lecturers confirmed their participation. Finally, 4 participants were selected based on the following inclusion criteria:

- Academic Qualifications: MA in TESOL, PhD in TESOL
- Subject matter expertise: Knowledge in the field of ELT and Curriculum Development
- Experience: Having more than 5 years of working experience in the ELT Department of HEIs
- Employment: Currently working as a lecturer in an ELT department
- Institutional diversity: Representing different universities

The primary purpose of this stage was to assess the consistency, flow and timing. The pilot study also revealed whether the questions were biased, leading, loaded, universal, lacking conceptual clarity, or asking for presumed expertise. This stage showed how questions might be comprehended and interpreted, how probes retrieved information, and the appropriate length of questions.

Table 6. Pilot Study Participants

Pilot study Participants	Qualifications/Position	Experience in working in the ELT	University. Department	Field of expertise
Lecturer A	MA in TESOL, PhD Candidate, Senior Lecturer	10 years	Kimyo International University, Uzbekistan, English Education	ELT
Lecturer B	MA in TESOL, Lecturer	8 years	Kimyo International University, Uzbekistan, English Education	ELT
Lecturer C	MA in TESOL, Senior Lecturer	6 years	Central Asian University, Uzbekistan	ELT
Lecturer D	MA in TESOL, Lecturer	6 years	Central Asian University, Uzbekistan	ELT

The research employs instrumental case study design (Yin, 2014), and pilot study serving to test the methods and procedures intended for use in the main qualitative study. Following the interview protocol, I conducted an in-depth semi-structured interview with participants, transcribed the interview and completed a reflexivity log to track emotions, bias and preconceptions (Lincoln & Guba, 1985). I position myself as an insider researcher in this study, as I belong to the HE context, being a lecturer in the ELT department. The reflexivity log contained two sections where I addressed my personal and interpersonal experiences before, during and after the interviews. To avoid researcher bias and help me reshape my research practice (Koopman et al., 2020), after each pilot interview, I recorded my responses to the questions: “How are my unique perspectives influencing the interview process?”, “How does my professional and personal background shape my interest in this topic?” (Olmos-Vega et al., 2018). Regarding interpersonal reflexivity, I addressed these questions: “What relationships exist and how do they influence the interview and the participants involved?”, “How are power dynamics operating for myself and pilot study participants?” (Olmos-Vega et al., 2018; Walsh, 2003). This reflexivity process helped me analyze my position and provided useful insights into participants’ responses, revealing “implicit biases in the researcher’s approach” (Finley, 2017).

Before data analysis, the transcriptions were sent to the interview participants for member checking (Creswell, 2018). Member-checking is a process where the research participants are involved in the research process by reviewing data transcripts and providing feedback (Birt et al., 2016; Varpio et al., 2017). The participants reviewed the transcripts and confirmed the accuracy of the information. However, Lecturer A and Lecturer D requested some changes. Lecturer A asked for the removal of information related to the university policy due to sensitivity concerns, while Lecturer D wanted to add some examples to the question about the important components of the ELT curriculum. As a result, the researcher decided to include an additional follow-up question to Section A (Curriculum Experience), asking participants to provide some examples or cases to support their answers.

2.2. Ethical Standards of the Pilot Study

This pilot study received ethical approval from the Central Asian University Institutional Review Board (IBR) under approval number CAUIRB2025057 granted on 17 September, 2025. The participants of the pilot study were presented with an informed consent form. The form includes a brief introduction to the research, the benefits to the subject and investigator, and some possible risks. The consent form also informs participants about their right to withdraw from this study at any time, and that all the information they provide will remain private and confidential (Creswell & Poth, 2018). Along with the consent form, the participants were also provided with a document that contains the information about the purpose and procedure of the research including the research questions, aims of the pilot study stage,

and time frame. The list of possible topics of the interview questions were also introduced in the document.

Regarding data storage and retention, to protect the participants from embarrassment, harm or stress, anonymization is a significant part of this pilot study (Creswell, 2018; Thomas, 2021). All the data of the pilot study, including consent forms are stored and managed according to the best practices for managing qualitative data (Corti et al., 2020). The transcripts are stored into a separate folder that is protected with a password, and the names of the participants are pseudonymized and used as the file names (Creswell, 2018). The drafts and improved versions of interview protocols, video recordings, and transcriptions of pilot interviews are also stored in a password-protected (Creswell, 2018) files, put in chronologically organized folders named by numeral identifier (Neale, 2021; Saldana, 2003). The data of the pilot study will be retained for a minimum of 6 years after publication (UK Data Service, 2023). This period of retention enables the verification of findings, possible secondary analysis, being complied with research integrity standards (Resnik et al., 2015). Researchers interested in accessing the documents might contact the author with a formal request which describes the purpose and objectives of their research.

3. Results of Modification

3.1. Refinements After External Review

The interview protocol underwent three major refinements which happened after external review by expert researchers, internal review by private university lecturers, and after pilot study. In terms of clarity of wording and language appropriateness, the experts recommend shortening the lengthy questions (*e.g., What courses, teaching methods, approaches, and techniques do you consider the most useful that should be integrated into the ELT curricula?*), break down complex questions into two separate questions, and simplify terminology (Table 6), as long questions can overwhelm or confuse the interviewees (Bryman, 2012). One of the experts commented on the topics covered in the questions related to curriculum development and evaluation that it lacks questions that ask the elements of the curriculum such as learning outcomes, objectives and syllabi. The reviewers also suggested modifying the tone of some questions as they sound like evaluating performance of lecturers (*e.g., What formal or informal methods do you use to evaluate the courses you teach?*) rather than asking for their insights. It was suggested to link the question back to the curriculum evaluation by adding “...and how does this relate to evaluating the curriculum as a whole?”. Another important suggestion was related to the sensitivity of some questions that ask what the participants would like to change in the current curriculum or share certain weaknesses of the curriculum. This type of questions requires verbal or written assurance of confidentiality (Creswell, 2018). Reviewers also recommended including a debriefing form that outlines objectives of the research and confidentiality matters and contact details of the researcher.

Table 7. Modification of Complex Questions After Expert Review

Stage 2	Original Questions	Expert Reviewer Comment	Revised version
Expert Review	<i>What courses, teaching methods, approaches, and techniques do you consider the most useful that should be integrated into the ELT curricula?</i>	There are many elements in one question	<i>What courses or teaching methods do you consider most useful for preparing future teachers?</i> <i>Could you explain why?</i>

Expert Review	<i>How would you describe the balance between theoretical and practical experience (teaching skills) that should be taught in the ELT curriculum?</i>	This question is too long	<i>How well does the curriculum balance theoretical knowledge and practical teaching skills?</i>
Expert Review	<i>What types of assessment are implemented? Could you provide specific examples?</i>	Some questions contain multiple questions inside one, which reduces clarity.	a) <i>What types of assessment are used in the curriculum?</i> b) <i>Could you share a few examples?</i>
Expert Review	<i>What assessment methods should be included in the ELT curriculum to fully equip pre-service teachers with the most essential teaching skills?</i>	To maintain clarity and prevent leading the respondent, break this into two separate questions. Divide the prompt into two clear and sequential questions. The first discusses <i>how</i> (assessment methods), while the second addresses <i>what</i> (essential skills)	<i>What types of assessment are used in the curriculum?</i> <i>Can you share a few examples?</i>

3.2. Refinements After Internal Review

The most significant modifications occurred after the internal review stage with the lecturers who work at private universities. The opening questions 2 and 3 (e.g., *What do you like the most about this program? What would you like to change in this program? If you have a chance to redesign the current curriculum, what would you change?*), which asked about strength of the program and possible modifications that might be suggested by the participants were criticized for not being appropriate warm-up questions, as they were a bit challenging and required to use problem solving skills right from the beginning. The review also revealed that most questions in the section B (core questions) subsections ii and iii need clarification prompts and examples (e.g., *How should an effective ELT curriculum address specific needs of pre-service teachers in terms of language and professional preparation?*). Some words like “feature”, “elements”, “components”, “structure”, “effective” were suggested to be specified as they were considered ambiguous.

Table 8. Modification of Questions After Internal Review

Stage 3	Original Questions	Expert Reviewer Comment	Revised version
Internal Review	<i>What do you like the most about this program? What would you like to change in this program?/ If you have a chance to redesign the current curriculum, what would you change?</i>	To me, the question does not sound very light. I mean is it generic? Is it a light and warm-up question?	<i>How would you describe the unique characteristics of ELT education in private universities?</i>
Internal Review	<i>How do these key elements contribute to the effectiveness of the ELT curriculum?</i>	What does 'effective' mean here? I suggest making it more precise.	<i>What does “effective curriculum” mean to you?</i>

Internal Review	<i>To make the curriculum more effective, what courses do you consider the most useful for preparing future English language teachers?</i>	I wonder what your expected answer here would be. I think obviously the respondent will say they are essential etc. Do you think you could also clarify which ones are more impactful to the success of a program?	<i>Which ones are more impactful to the success of a program, or arrange them in order of importance?</i>
Internal Review	<i>What are the most important features of an ELT curriculum that prepare pre-service ELT teachers effectively?</i>	For me the term 'feature' is somewhat ambiguous. Would you clarify what kind of features? Or in relation to what?	<i>What essential components should be included in an effective ELT teacher preparation curriculum?</i>

3.3. Refinements After Pilot Study

The piloting stage of the study simulated the interview stage of the actual research. The time frame of the interviews ranged between 50-70 minutes. The appropriateness and clarity of the questions were tested and improved. However, it was noticed that some questions required cuing (hinting) for certain participants who experienced cognitive load and memory delays (Foddy & Foddy, 1993). Another important finding was that some questions or prompts elicited bounded responses while some required unbounded. In other words, some questions generated limited and narrow responses (bounded) while some had unbounded extensive responses (Weller et al., 2018). This identification helped modify some questions to maintain the balance to avoid discussions that are either too narrow or too broad.

Pilot study also indicated that some questions should be revised to be more personalized and informal to foster a closer, and better relationship with participants and reduce stress. For example, the question “What courses are essential or should be a part of the curriculum?” was changed to “If you were an ELT curriculum developer, what courses would you add or remove from the curriculum?”. During the interviews, it was found that some participants were not familiar with the content and structure of the curriculum as they were not involved in the development process. Therefore, curriculum related questions were replaced with other follow-up questions focusing on the courses they teach. Some phrases such as “working with curriculum” were changed to “working in the ELT program” or “working with pre-service teachers”, as some participants highlighted that they are not directly involved in curriculum development. Finally, considering that lecturers typically work with syllabi, some follow-up questions about the syllabi were added before moving the discussion on curriculum.

Table 8. Modification of the Drafts

Stage	Draft	Number of items revised
External Review	Draft 2	7
Internal Review	Draft 3	10
Pilot study	Draft 4	18

During the interview some participants produced overly diverse answers, which led to not being able to have adequate discussion to other questions within the given time. This stage also helped reduce the interviewer’s stress and provided an opportunity for training before the actual interview. Furthermore,

this stage confirmed the benefits of piloting regarding the evaluation of practicality of research procedures.

4. Conclusion

Qualitative research should be a cyclical, interactive process which prioritizes decisions influenced by the research aims and “the voice of participants” (Demirci, 2024). Ensuring rigor in qualitative inquiry demands trust and rapport, in-depth discussions, being unbiased “to make the participants the stars of the show” (Qu & Dumay, 2011). This qualitative study puts lecturers at the center of the ELT program, recognizing their role as key contributors to the enhancement of curriculum development and evaluation from the bottom-up. The preliminary phase of the research described in this paper demonstrates how lecturers could be involved in the entire research procedure not only as participants, but also as reviewers who could improve the instrument of the study. Moreover, ELT program lecturers could provide some insights on the current state of ELT curriculum and how it might be improved. They are insiders who could reveal valuable information on curriculum evaluation based on their pedagogical expertise and needs of their students. They also possess “theoretical sensitivity” (Maxwell, 2013) of curriculum development and evaluation approaches, expertise on theoretical principles and practical constraints. Considering the data gained from the participants who are the core stakeholder of ELT curriculum, curriculum experts and developers are expected to contribute to the development of curriculum evaluation framework and guidelines which are adjusted to the context of ELT program of private universities.

This multi-step interview protocol refinement process enabled the systematic development of the robust qualitative research instruments. Each step was thoroughly documented, maintaining transferability and transparency of the study. The protocol underwent four rounds of review and refined after completing each stage, which facilitated in-depth exploration of the research topic. The comments of the external and internal reviewers provided an opportunity “to fine-tune the interview protocol” (Castillo-Montoya, 2016), while pilot study helped the researcher to witness how the interview protocol functions in reality prior to conducting actual interview.

References

- Bezborodova, A. (2023). *Language practices and processes in English-medium higher education in Uzbekistan* (Doctoral dissertation, University of Westminster, London, United Kingdom). <https://doi.org/10.34737/w4642>
- Bezborodova, A., & Radjabzade, S. (2022). English in higher education in the Kyrgyz Republic, Tajikistan, and Uzbekistan. *World Englishes*, 41(1), 72–91.
- Bickman, L., & Rog, D. J. (Eds.). (2009). *The Sage handbook of applied research design* (2nd ed.). Sage.
- Braaten, B., Kramer, A., Henderson, E., Kajfez, R., & Dringenberg, E. (2020). Accessing complex constructs: Refining an interview protocol. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE.
- Bryman, A. (2012). *Social research methods*. Oxford University Press
- Carl, A. (2009). *Teacher empowerment through curriculum development theory into practice*. Juta & Company Ltd.
- Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report*, 21(5), 811–831.
- Cheema, K. U. R. (2021). Barriers to curriculum implementation: A systematic literature review. *European Journal of Education Studies*, 8(11), 318-339. <https://doi.org/10.46827/ejes.v8i11.4001>
- Coker, A. D., & Akande, O. (2025). Preparation, development, and refinement of the interview protocol in qualitative research. *The Qualitative Report*, 30(8), 4205-4235. <https://doi.org/10.46743/2160-3715/2025.7964>
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2020). *Managing and sharing research data: A guide to good practice* (2nd ed.). Sage.
- Creswell, J. W. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approach*. US: Sage.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications, Inc.
- Cohen, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12(3), 229–238.
- Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*, 17(1), 7–28.
- Demirci, J. R. (2024). About research: Conducting better qualitative interviews. *MCN: The American Journal of Maternal/Child Nursing*, 40(1). <https://doi.org/10.1177/08903344231213651>
- Erasmus+ National Office in Uzbekistan. (2024, October). *Overview: Higher education profile (Uzbekistan)*. <https://erasmus.uz/en/page/11-overview>
- Finlay, L. (2017). Championing ‘reflexivities’. *Qualitative Psychology*, 4(2), 120–125. <https://doi.org/10.1037/qup0000075>

- Foddy, W., & Foddy, W. H. (1993). *Constructing questions for interviews and questionnaires: Theory and practice in social research*. Cambridge University Press.
- Fullan, M. (1991). *The meaning of educational change*. Teacher College Press.
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2013). *The basic guide to supervision and instructional leadership* (3rd ed., pp. 250-262). Pearson.
- Hafeez, M., Kazmi, Q. A., Tahira, F., Hussain, M. Z., Ahmad, S., Yasmeen, A., & Saqi, M. I. (2021). Challenges in curriculum implementation at higher education level. *International Journal of Learning and Teaching*, 7(1), 30-38. <https://doi.org/10.18178/ijlt.7.1.30-38>
- Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2007). The new instructional leadership: Creating data driven instructional systems in school. *Journal of School Leadership*, 17(2), 159–194.
- Handler, B. (2010). Teacher as curriculum leader: A consideration of the appropriateness of that role assignment to classroom-based practitioners. *International Journal of Teacher Leadership*, 3, 32–42.
- Glatthorn, A. A., Boschee, F., & Whitehead, B. M. (2016). *Curriculum leadership: Strategies for development and implementation*. SAGE Publications, Inc.
- Gordon, W., Taylor, R., & Oliva, P. (2019). *Developing the Curriculum*. Pearson Education.
- Ismail, N., Kinchin, G., & Edwards, J. (2018). Pilot study, does it really matter? Learning lessons from conducting a pilot study for a qualitative PhD thesis. *International Journal of Social Science Research*, 6(1), 1–17.
- Janghorban, R., Latifnejad Roudsari, R., & Taghipour, A. (2014). Pilot study in qualitative research: The roles and values. *Hayat*, 19(4), 1–5.
- Jones, S. R., Torres, V., & Arminio, J. (2013). *Negotiating the complexities of qualitative research in higher education: Fundamental elements and issues*. Routledge.
- Kärkkäinen, K. (2012). Bringing about curriculum innovations. *OECD Education Working Papers*, No. 82. OECD Publishing. <https://doi.org/10.1787/5k95qw8xzl8s-en>
- Kim, Y. (2010). The pilot study in qualitative inquiry: Identifying issues and learning lessons for culturally competent research. *Qualitative Social Work*, 10(2), 190–206. <https://doi.org/10.1177/1473325010362001>
- Khan, M. A., Khan, S. A., & Khan, M. A. (2023). Challenges and opportunities in curriculum development and implementation in higher education: A systematic review. *Advances in Medical Education and Practice*, 14, 123-134. <https://doi.org/10.2147/AMEP.S395029>
- Koopman, W. J., Watling, C. J., & Ladonna, K. A. (2020). Autoethnography as a strategy for engaging in reflexivity. *Global Qualitative Nursing Research*, 7, 2333393620970508. <https://doi.org/10.1177/2333393620970508>
- Kvale, S. (2007). *Doing interviews*. Sage Publications.
- Louvel, S. (2013). Understanding change in higher education as bricolage: How academics engage in curriculum change. *Higher Education*, 66(6), 669–691. <https://doi.org/10.1007/s10734-013-9628-6>
- Lynch, B.K. (2003). *Language assessment and programme evaluation*. Edinburgh: Edinburgh University Press.

Mandel, D. (2015). *Uncovering the institutions of accountability policy: Capturing the institutions of policy makers and principals*. University of Arizona.

Maxwell, J. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Sage.

McGrath, C., Barman, L., Stenfors-Hayes, T., Roxå, T., Silén, C., & Bolander Laksov, K. (2016). The ebb and flow of educational change: Change agents as negotiators of change. *Teaching and Learning Inquiry*, 4(1), 1–15. <https://doi.org/10.20343/teachlearninqu.4.2.9>

Micallef, B. (2025a). The role of lecturers as agents of change in higher education. *Frontiers in Education*, 10, 1580611.

Micallef, B. (2025b). Enhancing curriculum relevance and student engagement: The role of lecturers as agents of change in higher education. In *Proceedings of the 9th World Conference on Qualitative Research*. Galoá.

Merriam, S. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Neale, B. (2021). *The Craft of Qualitative Longitudinal Research*. Sage.

Nevenglosky, E. A., Cale, C., & Aguilar, S. P. (2019). Barriers to effective curriculum implementation. *Research in Higher Education Journal*, 36(1), 31-48.

Nieveen, N. (2011, September 14). Teachers' professional development in curriculum design in the Netherlands. Paper presented at the European Conference for Educational Research, Berlin.

Olmos-Vega, F. M., Dolmans, D. H. J. M., Guzman-Quintero, C., Stalmeijer, R. E., & Teunissen, P. W. (2018). Unravelling residents' and supervisors' workplace interactions: An intersubjectivity study. *Medical Education*, 52(7), 725–735. <https://doi.org/10.1111/medu.13600>

Peacock, M. (2009). The Evaluation of Foreign-language-teacher Education Programmes. *Language Teaching Research*, 13(3), 259-278.

Priestley, M. (2011). Whatever happened to curriculum theory? Critical realism and curriculum change. *Pedagogy, Culture and Society*, 19(2), 221–238.

Priestley, M., Robinson, S., & Biesta, G. J. J. (2012). Teacher agency, performativity and curriculum change: Reinventing the teacher in the Scottish Curriculum for Excellence? In B. Jeffrey & G. Troman (Eds.), *Performativity across UK education: Ethnographic cases of its effects, agency and reconstructions* (pp. 87–108). E&E Publishing.

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage.

Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting & Management*, 8(3), 238-264. <https://doi.org/10.1108/11766091111162070>

Rasulov, A. (2024). Reforms in the higher education system of Uzbekistan aimed at preparing competitive personnel. Embassy of Uzbekistan in Ukraine. <https://www.uzbekistan.org.ua/en/145-biblioteka/7214-reforms-in-the-higher-education-system-of-uzbekistan-aimed-at-preparing-competitive-personnel.html>

Rea-Dickins, P. and Germaine, K.P. (1998). The price of everything and the value of nothing: trends in language programme evaluation. In Rea-Dickins, P. and Germaine, K.P. (Eds.), *Managing evaluation and innovation in language teaching: building bridges* (pp. 3–19). London: Longman.

Resnik, D. B., Morwitz, V. G., & Hesse, B. W. (2015). Survey research and scientific integrity: What do researchers think? *Research Integrity and Peer Review*, 1, Article 3. <https://doi.org/10.1080/08989621.2017.1327813>

Ruziev, K., & Burkhanov, U. (2018). Uzbekistan: Higher education reforms and the changing landscape since independence. 25 Years of Transformations of Higher Education Systems in Post-Soviet Countries: Reform and Continuity, 435–459.

Saldana, J. (2003). *Longitudinal qualitative research: Analyzing change through time*. Alta Mira Press.

Scorsolini-Comin, F. (2020). *Técnicas de entrevista: Método, planejamento e aplicações*. Vetor Editora.

Secomb, J. M., & Smith, C. (2011). A mixed method pilot study: The researchers' experiences. *Contemporary Nurse*, 39, 31–35.

Shadiev, R., Lin, L., Shen, S., & Yu, M. (2025). Review of research on design thinking in K-12 education. *Educational Research Review*.

Sieber, J. E. (1998). Planning ethically responsible research. In L. Bickman & D. J. Rog (Eds.), *Handbook of applied social research methods* (pp. 127–156). Sage.

Stephens, J. C., Hernandez, M. E., Román, M., Graham, A. C., & Scholz, R. W. (2008). Higher education as a change agent for sustainability in different cultures and contexts. *International Journal of Sustainability in Higher Education*, 9(3), 317–338. <https://doi.org/10.1108/14676370810885916>

Thomas, G. (2021). *How to do your case study*. Sage Publications.

UK Data Service. (2023). *Research data lifecycle*. <https://ukdataservice.ac.uk/learning-hub/research-data-management/>

UzDaily. (2025, May 29). *Over the past decade, the number of universities in Uzbekistan has increased 3.2 times — reaching 222 institutions*. UzDaily.com. <https://www.uzdaily.uz/en/over-the-past-decade-the-number-of-universities-in-uzbekistan-has-increased-32-times-reaching-222-institutions/>

Van der Heijden, H., Geldens, J., Beijaard, D., & Popeijus, H. (2015). Characteristics of teachers as change agents. *Teachers and Teaching: Theory and Practice*, 21(6), 681–699. <https://doi.org/10.1080/13540602.2015.1044328>

Van Wijk, E., & Harrison, T. (2013). Managing ethical problems in qualitative research involving vulnerable populations, using a pilot study. *International Journal of Qualitative Methods*, 12(1), 570–586. <https://doi.org/10.1177/160940691301200130>

Walsh, R. (2003). The methods of reflexivity. *Humanistic Psychologist*, 31(4), 51–66. <https://doi.org/10.1080/08873260309596005>

Weller, S. C., Vickers, B., Bernard, H. R., Blackburn, A. M., Borgatti, S., Gravlee, C. C., & Johnson, J. C. (2018). Open-ended interview questions and saturation. *PloS One*, 13(6), e0198606. <https://doi.org/10.1371/journal.pone.0198606>

World Bank. (2023). Implementation completion report (ICR) review: Modernizing higher education project (P128516). <https://documents.worldbank.org/curated/en/099050124112542784/pdf/P1285161216e9e09b1bcd21bf97679714b0.pdf>

Yang, W. (2009). Evaluation of teacher induction practices in a US university English language program: Towards useful evaluation. *Language Teaching Research*, 13(1), 77-98.

Yin, R. (2018). *Case study research and applications: design and methods*. Thousand Oaks, Sage Publications.

Zhao, S. (2011). Actors in language planning. In E. Hinkel (Ed), *Handbook of research in second language teaching and learning: Volume 2* (905-923). New York: Routledge.

Zhang, S., Abd Aziz, M. F. Asri, A. S., & Saidi, Z. A. (2025). Factors influencing female Lecturers' well-being in higher education: A comprehensive review. *Journal of Institutional Research South East Asia*, 23(1), 418-443.

Appendix A

Exploring Stakeholder Perspectives on Effective Curriculum Evaluation Methods for ELT Pre-Service Teacher Education at Private Universities in Uzbekistan

Semi-Structured Interview Questions for ELT Curriculum Evaluation Research in Uzbekistan

Introduction

I want to express my gratitude for agreeing to participate in this research. My name is _____. I am a Lecturer at the English Language Teaching and Educational Management (ELTEM) Department at _____ University in Tashkent and a PhD student at the faculty of _____, at _____ University.

The main purpose of this study is to explore and understand perceptions, experiences, and recommendations regarding the evaluation of English Language Teaching (ELT) curriculum in pre-service education in private higher education institutions in Uzbekistan. This research comprises two distinct stages. First, the study aims to identify the essential needs of ELT teachers by establishing criteria for an effective ELT curriculum based on perspectives from key stakeholders: university lecturers who closely work with the target population (pre-service teachers, their parents, and in-service teachers) and curriculum experts who have experience in developing ELT curricula. Based on the perceptions of ELT program lecturers and curriculum experts, essential components and approaches of the ELT curriculum will be identified. Second, the research focuses on developing guidelines and a framework (or modifying an existing framework) for evaluating ELT curricula, informed by the expertise and experiences of ELT lecturers and curriculum specialists.

You are going to participate in the interview. The interview is planned to take 60 to 90 minutes. I would like to ask your permission to audio (video) record the interview process to assure the accuracy of your thoughts and experiences. This will also help us engage in a more fluid conversation focusing on the interview in-depth, asking follow up questions instead of pausing to take notes.

Any information you share will be held confidential, secure and anonymous. Only the researcher will have access to the information and will aggregate the results from the all interviews, assuring you that the findings will not be attributed to any individual or any location.

Please know that your participation is entirely voluntary, and you may choose to stop or pause the interview at any time without any consequences. Your comfort and consent are very important throughout this process.

While responding to the interview questions, you are expected to provide detailed, honest answers based on your experience and expertise. Please feel free to share specific examples or situations that illustrate your points. Your reflections on both successes and challenges are valuable to deepen our understanding.

Are there any questions about the interview process?

Are you willing to participate in this interview?

TOPICS

- The sequence of courses throughout the English Language Teaching curriculum
- The importance of practical and theoretical balance
- The importance and need of core language components (grammar, vocabulary, receptive and productive skills)
- Pedagogical theory and teaching methodologies
- Integration of technology in language teaching
- Material design and adaptation
- Practical teaching skills and practicum experiences
- Language assessment and evaluation
- Curriculum development and evaluation

Interview Questions for University Lecturers

SECTION A

Opening questions: Building Rapport

1. Could you please introduce yourself?
2. What is your overall experience in working with the ELT curriculum/working as a lecturer in the ELT program?
3. What courses do you teach?
4. What do you like the most about working with pre-service ELT teachers?
5. How would you describe the unique characteristics of ELT education in private universities?/
What are some unique characteristics of ELT education in the context of a private university?

SECTION B

Core questions

i. Curriculum experience

Main questions	Prompts/Follow-up questions
1. a. What do you know about the current ELT curriculum you work with? b. Tell me how the ELT curriculum is organized in your university.	a. <i>What courses are currently part of the curriculum? Could you mention a few examples?</i> b. <i>How well does the curriculum balance theoretical knowledge and practical teaching skills?</i>

	c. <i>How is the practicum (teaching practice) organized?</i> <i>(When does it start and how long does it last?)</i> d. <i>What types of assessment are used in the curriculum? Can you share a few examples?</i> e. <i>How does the curriculum incorporate teaching technologies (e.g., digital tools, online platforms)?</i> f. <i>Can you give examples of how these are used in practice? (Are there any courses on instructional technology)</i> g. <i>Tell me more about the courses you teach.</i> h. <i>Tell me about the components of the syllabus. What are some essential criteria to design a syllabus?</i> i. <i>How often and why do you change the syllabus?</i>

ii. Perceptions of effective curriculum

Main questions	Prompts/Follow-up questions
2. a. How do you define “effective curriculum?”/What does the word “effective” mean to you? b. What key elements do you think should be part of an effective ELT curriculum?	a. <i>Which elements are more impactful to the success of a program, or arrange them in order of importance?</i> b. <i>To make the curriculum more effective, what courses do you consider the most useful for preparing pre-service ELT teachers?/anything you would like to add to make the curriculum more effective?</i> c. <i>If you were a curriculum developer, what courses would you add or remove?/If you were a curriculum developer, what changes would you make?</i> d. <i>What assessment methods do you use?</i> e. <i>What assessment methods do you find most effective for evaluating pre-service teachers’ readiness? Why?</i> f. <i>Based on your experience, what are the three to five most essential skills/knowledge that the ELT curriculum</i>

	<i>must ensure all pre-service teachers graduate with?</i>
3. How do you think the current ELT curriculum prepares pre-service teachers for their future careers?	<i>a. Could you describe a situation where graduate students were well-prepared or underprepared for actual teaching context?</i>
4. How do you think (your personal opinion) the ELT curricula differ between public and private universities (in terms of their practical focus, resource allocation, and responsiveness to market demands)?	<i>a. Can you provide specific examples of how these differences might impact student learning outcomes and career preparation?</i>

iii. Criteria and approaches for curriculum evaluation

Main questions	Prompts/Follow-up questions
5. How would you define “curriculum evaluation” in simple terms?	<i>a. What do you know about the evaluation stage of the current ELT curriculum?</i> <i>b. Have you personally participated and/or observed a curriculum evaluation process?</i> <i>c. If yes, can you share your experience of how it was conducted (stages)?</i> <i>d. Who was involved in the curriculum evaluation process?</i>
6. a. What should be evaluated in the ELT curriculum in your opinion? b. If we think about the list of criteria for curriculum evaluation, what would you include?./If you were asked to evaluate the ELT curriculum and make a list of criteria, what would you include?	<i>a. Could you suggest criteria/qualities that are essential for evaluating ELT curriculum in a private university context?</i> <i>b. Who should be involved in the curriculum evaluation process?</i> <i>c. What about your course evaluation? tell me more about the process</i> <i>d. How do you know the course you have taught was successful?</i> <i>e. What formal or informal methods do you use to evaluate the courses you teach?</i>

	f. <i>How does this relate to evaluating the curriculum as a whole?</i> g. <i>Can you share a specific case where student outcomes clearly showed the strengths and weaknesses of the curriculum?</i>
7. a. How should the ELT curriculum evaluation process be conducted in the context of a private university? b. If we put curriculum evaluation into specific stages, how can it be effectively organized?	a. <i>How can curriculum evaluation be effectively integrated to ensure continuous improvement?</i> b. <i>How can curriculum evaluation demonstrate alignment with educational goals/objectives?</i> c. <i>How do you determine whether the curriculum is meeting the needs of pre-service teachers (any approaches?)</i> d. <i>In your opinion, what are the most effective ways (approaches) to conduct the ELT curriculum evaluation?</i> e. <i>What are the steps? Why are these approaches effective?</i> f. <i>How feedback from evaluation should be used to improve curriculum?</i> g. <i>How frequently should the ELT curriculum be evaluated and updated? Why?</i>

SECTION C

Closing questions

1. What barriers exist to implementing effective curriculum evaluation in a private university? How would you tackle them?
2. Based on our discussion today, what do you believe is the single most important factor that should be considered when evaluating an ELT curriculum for pre-service teachers?
3. What final advice would you give to someone developing an evaluation framework for ELT pre-service teacher education programs in Uzbekistan?
4. Do you have anything else to share about ELT curriculum content and development in the context of private university?

Generic probing questions

1. Could you provide a specific example? / Can you give an example on what you have just mentioned?
2. Could you elaborate on this/that point?

3. Why do you believe this approach/method/course was/wasn't effective?
4. How could that be improved?
5. What do you mean by.....?
6. How is that different/similar from your expectations?

Appendix B

Final Alignment Matrix

Research Questions	Core Questions	Prompts & follow-ups	Rationale
RQ1 How do university lecturers and curriculum experts perceive an effective ELT curriculum at private universities in Uzbekistan?	1. What key elements (e.g., courses, practicum, assessment) do you believe should be part of an effective ELT curriculum? (to meet the needs of pre-service teachers)	a. <i>What courses are currently part of the curriculum? Could you mention a few examples?</i> b. <i>How well does the curriculum balance theoretical knowledge and practical teaching skills?</i> c. <i>How is the practicum (teaching practice) organized?</i> <i>(When does it start and how long does it last?)</i> d. <i>What types of assessment are used in the curriculum? Can you share a few examples?</i> e. <i>How does the curriculum incorporate teaching technologies (e.g., digital tools, online platforms)?</i> f. <i>Can you give examples of how these are used in practice? (Are there any courses on instructional technology)</i> g. <i>Tell me more about the courses you teach.</i> h. <i>Tell me about the components of the syllabus. What are some essential criteria to design a syllabus?</i>	The questions progress from general to specific, highlighting what elements should be inside the curriculum. The first follow-up question encourages obtaining broad and diverse perspectives of participants. Next questions are more specific, asking about the necessary courses, assessment, and skills that should be integrated inside the ELT curriculum that prepares pre-service teachers for real teaching context

		<i>i. How often and why do you change the syllabus?</i>	
RQ2 What criteria do university lecturers and curriculum experts perceive as essential for effective ELT curriculum evaluation in private universities in Uzbekistan?	2. a. How do you define “effective curriculum?”/What does the word “effective” mean to you? b. What key elements do you think should be part of an effective ELT curriculum? 3. How do you think the current ELT curriculum prepares pre-service teachers for their future careers? 4. How do you think (your personal opinion) the ELT curricula differ between public and private universities (in terms of their practical focus, resource allocation, and responsiveness to market demands)?	a. <i>Which elements are more impactful to the success of a program, or arrange them in order of importance?</i> b. <i>To make the curriculum more effective, what courses do you consider the most useful for preparing pre-service ELT teachers?/anything you would like to add to make the curriculum more effective?</i> c. <i>If you were a curriculum developer, what courses would you add or remove?/If you were a curriculum developer, what changes would you make?</i> d. <i>What assessment methods do you use?</i> e. <i>What assessment methods do you find most effective for evaluating pre-service teachers’ readiness? Why?</i> f. <i>Based on your experience, what are the three to five most essential skills/knowledge that the ELT curriculum must ensure all pre-service teachers graduate with?</i>	The purpose of these questions is to identify the perceptions of ELT lecturers in terms of curriculum effectiveness, the courses, assessment methods, the balance between theory and practice that they consider important. The questions help explore general criteria of effective ELT curriculum
RQ3 What approaches do university lecturers and curriculum experts recommend for effective	5. How would you define “curriculum evaluation” in simple terms?	a. <i>What do you know about the evaluation stage of the current ELT curriculum?</i> b. <i>Have you personally participated and/or observed a</i>	The main focus of these questions is to identify how the ELT curriculum should be organized, what

curriculum evaluation at private universities in Uzbekistan?		<i>curriculum evaluation process?</i> <i>c. If yes, can you share your experience of how it was conducted (stages)?</i> <i>d. Who was involved in the curriculum evaluation process?</i>	could be the stages and approaches for evaluations from the perspective of lecturers. The questions also focus on the other stakeholders that should participate in the evaluation process and the frequency of evaluation. The questions also cover how lecturers evaluate their own courses
	6. a. What should be evaluated in the ELT curriculum in your opinion? b. If we think about the list of criteria for curriculum evaluation, what would you include?./If you were asked to evaluate the ELT curriculum and make a list of criteria, what would you include?	a. <i>Could you suggest criteria/qualities that are essential for evaluating ELT curriculum in a private university context?</i> b. <i>Who should be involved in the curriculum evaluation process?</i> c. <i>What about your course evaluation? tell me more about the process</i> d. <i>How do you know the course you have taught was successful?</i> e. <i>What formal or informal methods do you use to evaluate the courses you teach?</i> f. <i>How does this relate to evaluating the curriculum as a whole?</i> g. <i>Can you share a specific case where student outcomes clearly showed the strengths and weaknesses of the curriculum?</i>	

	7. a. How should the ELT curriculum evaluation process be conducted in the context of a private university? b. If we put curriculum evaluation into specific stages, how can it be effectively organized?	a. <i>How can curriculum evaluation be effectively integrated to ensure continuous improvement?</i> b. <i>How can curriculum evaluation demonstrate alignment with educational goals/objectives?</i> c. <i>How do you determine whether the curriculum is meeting the needs of pre-service teachers (any approaches?)</i> d. <i>In your opinion, what are the most effective ways (approaches) to conduct the ELT curriculum evaluation?</i> e. <i>What are the steps? Why are these approaches effective?</i> f. <i>How feedback from evaluation should be used to improve curriculum?</i> g. <i>How frequently should the ELT curriculum be evaluated and updated? Why?</i>	

Appendix C

Feedback Summary

Criteria	Reviewer 1	Reviewer 2	Reviewer 3
Clarity of wording	Some questions are quite long or contain multiple sub-	Key to a curriculum are syllabus, methodology, and assessment. Then,	Some main questions and probes are overly lengthy or seek to cover numerous subjects simultaneously.

	questions within a single item, which may overwhelm participants or reduce the clarity of what is being asked. A few terms, such as “effective curriculum,” “components,” “innovative technologies,” or “curriculum evaluation,” are somewhat broad and may benefit from brief clarification or examples to ensure consistent interpretation across participants.	objectives, teacher role, student role, materials, and resources. Some of these are included in the first question but not all. I think this should be added. This question should be revised for lecturers: What do you think, what should be considered in the ELT curriculum evaluation?	Divide the prompt into two clear and sequential questions. The first discusses <i>how</i> (assessment methods), while the second addresses <i>what</i> (essential skills). Simplify the question by utilizing the core, umbrella term ("methods"), and then use the given Generic Probing Questions for gathering details about "approaches and techniques." You should then use the existing generic probes
Relevance		quite inclusive	e questions are clearly and directly relevant to the three main research questions (RQs). The protocols successfully provide the necessary path to obtain perceptions of effectiveness (RQ1), evaluation criteria (RQ2), and effective approaches/recommendations (RQ3). The interview is structured into three sections: descriptive context (Section B.i), perceptions (B.ii), and evaluation (B.iii), ensuring that all RQs are covered systematically. The material is ideal for gathering data to

			directly inform the final recommended evaluation framework.
Coverage	Overlap with lecturers' questions Many questions are identical; consider designing distinct roles: Lecturers: focus on implementation and student outcomes Curriculum experts: focus on design, alignment, evaluation frameworks, policy	Key to a curriculum are syllabus, methodology, and assessment. Then, objectives, teacher role, student role, materials, and resources. Some of these are included in the first question but not all.	The protocols are comprehensive in terms of core RQs, however there is a significant stakeholder gap. While the thesis title and introduction reference a wide range of stakeholders (pre-service teachers, parents, and in-service teachers) whose needs should be considered while developing the curriculum, the current questions mainly address the perspectives of lecturers and experts. This omission puts the final evaluation framework at risk of being disconnected from the perspectives of the end users. A critical question must be included to both protocols in Section B(iii). In the Lecturer Protocol, Q2 and its subsequent prompts (a-e) are slightly redundant. To ensure a maximum 60-minute period, combine these into a single main question and then probe generically.
Sequencing and flow	Some questions are too leading or prescriptive <i>Example:</i> What teaching methods... should be included? What assessment	Yes, quite fine.	The general flow of the interview is great and logical. The protocol develops seamlessly from simple interpersonal inquiries (Section A) to descriptive questions about the current program (Section B.i), followed by more complicated and opinion-based questions about effectiveness (B.ii) and

	methods should be included? These imply that methods must be included. Paraphrase		evaluation recommendations (B.iii). This framework is good for qualitative research because it allows participants to build comfort and context before delving into the most abstract and prescriptive issues, resulting in the collection of rich data.
Language appropriateness	Too many prompts Some questions have 5–7 prompts, which will overwhelm participants. Consider grouping prompts into one combined question with optional follow-ups. Some questions require contextual clarification	Yes, readable and comprehensible	The language chosen is completely appropriate for the targeted participants' academic and professional backgrounds (university lecturers and curriculum experts). The distinction between the two protocols is also nicely performed, with the Expert guide correctly using more abstract, system-level vocabulary, whereas the Lecturer guide concentrates on practical program realities. However, you should be prepared to verbally clarify complex academic terms (such as "systematic approaches") to guarantee mutual understanding during the interview.
Ethical considerations	Your introduction is already strong. Minor improvements: Add a sentence about right to withdraw: <i>You may withdraw at any time</i>		The ethical forms given (Recruitment, Consent, Data Handling) should meet normal research ethics board requirements (e.g., use of pseudonyms, password security, data destruction), as it was previously mentioned. Questions needing criticism, such as Lecturer Protocol, Section A, Q4 (<i>"What would you like to change in this</i>

	<i>without consequences.</i> Add confidentiality clarification: <i>Data will be anonymized and stored securely following UPM ethical guidelines.</i>		<i>program?"),</i> have a low level of sensitivity, especially in a private HEI where job security might be a concern. Prior to asking this type of questions, you should orally reinforce the confidentiality guarantee. A debriefing form is required. This form should briefly clarify the study's goal, reaffirm confidentiality, provide contact information, and express the researcher's willingness to disclose aggregate findings. This grants the participants the right to be debriefed.
Overall suitability	Some questions are repetitive or overlap in meaning, particularly those related to curriculum components, course value, and theoretical–practical balance. Missing technical curriculum evaluation terminology Missing question on challenges	I think it covers what is needed for curriculum evaluation should the interviewer engage well with the participants. This engagement definitely requires wrestling with culture as well, which is essential for curriculum design but I don't know to what extent culture is the focus of the study. If it is, I think it can be added to the prompts I mentioned above.	The protocols are comprehensive, relevant, and well-designed for a qualitative study. They aim to obtain the depth of experience and detailed recommendations required to create the final evaluation framework, thereby meeting the study objectives.
Additional suggestions	Missing question on challenges		Given the current length and complexity of the interview, the candidate must prepare for conducting it to ensure it fits within the maximum 60-

	Add: “What challenges or barriers do you see in evaluating ELT curricula in private universities?” This will enrich your findings. Add: <i>What evaluation models or frameworks do you consider most appropriate for ELT curriculum evaluation in Uzbekistan?</i> Missing international benchmarking You have one question at the end, which is good. Add follow-up: Which countries’ ELT curriculum evaluation practices could serve as a model for Uzbekistan?		minute timeframe. All recruitment and consent documents must be completed in a professional academic tone, with all bracketed placeholders (e.g., [duration], [Your Name]) filled with final, specific information.

Appendix D

Original question	Revised Question	Interview Excerpt
What should be evaluated in the ELT curriculum? <i>(This question is too broad and sounds like interviewer checks interviewees knowledge which makes an interview process a bit stressful)</i>	If you were asked to evaluate the ELT curriculum and make a list of criteria, what would you include? <i>(This question is more structural and guided. It encourages interviewees to put themselves in the role of an evaluator and think practically what criteria they would use. The tone of question is more personal and task-oriented.)</i>	Okay, um, maybe it would be the alignment again, what is the framework that it is adopting or adapting, using yes and where it is reflected, how it is reflected. Another important criteria would be learning objectives, or outcomes, as we call them, right? And these learning outcomes, are they aligned with, with the content, with the assessment, also definitely it would be if it was the case with my current workplace, alignment with graduate attributes, because we have the attributes and throughout their studies, within any course, in any year, every learning outcome should reflect The graduate attributes. So that alignment would definitely be there. And of course, we would take a look at the assessment. Does it does it follow assessment principles? Is it reliable? Is it valid? I don't know. Is practicality in place? And also nowadays we're looking into SDGs. It's become very common, especially since our university has started adopting them and joined this initiative, right? So we are trying to make sure that we adhere to SDG four, at least, if not SDG seven. So what else? Yes, we also take a look at the types of the sources that were used, references that were used. So are the authors of the curriculum inspired by existing courses? So in

		what way is it different from that course? If it is Western context,

Appendix E

Approval letter from Institutional Review Board



CAU
CENTRAL
ASIAN
UNIVERSITY

CENTRAL ASIAN UNIVERSITY LLC
264 Milliy Bog street, Mirzo Ulugbek District,
Tashkent, Uzbekistan
+998 71 200 0522
info@centralasian.uz

Institutional Review Board (IRB)
University Ethics Committee

Date: 17/09/2025
Ref No: IRB/CAU/2025/057
Ethical Approval Code: CAU-IRB-2025-056

To Whom It May Concern,

The **Institutional Review Board (IRB) of Central Asian University** has reviewed the research proposal and survey instrument for the study entitled:

“Refining Qualitative Research Instruments through Participant Involvement: An Interview Protocol Refinement Process for ELT Curriculum Research in Uzbekistan”

Principal Investigator: Nozimakhon Uralova

RB Decision: Approved (Minimal Risk)

Following a comprehensive ethical review, the IRB confirms that the study complies with institutional ethical standards and international guidelines, including the **Declaration of Helsinki (2013 revision)**. The study adopts a qualitative research design focused on refining the interview protocol and employs semi-structured interviews, pilot studies, and expert reviews with university lecturers in Uzbekistan.

Ethical Review Summary

The research involves voluntary participation of adult university lecturers who contribute to the refinement of qualitative research instruments through interviews, reviews, and pilot study participation. The study does not involve vulnerable populations, clinical procedures, or sensitive personal data. Participants are fully informed about the purpose of the research and provide informed consent prior to participation. Participation is voluntary, and participants retain the right to withdraw at any stage without consequence. Data will be collected through anonymized interviews, and confidentiality is strictly maintained. Member checking procedures are implemented to ensure data accuracy and participant validation. The study poses minimal risk,



CENTRAL ASIAN UNIVERSITY LLC

264 Milliy Bog street, Mirzo Ulugbek District,
Tashkent, Uzbekistan
+998 71 200 0522
info@centralasian.uz

limited to potential professional reflection or mild cognitive engagement during interviews, and adheres to strict ethical standards for qualitative research.

Ethical Considerations Confirmed

- Informed consent was obtained prior to participation
- Voluntary participation with the right to withdraw at any time
- Confidentiality and anonymity are ensured in all data handling
- Use of member checking to enhance credibility and participant validation
- No collection of sensitive personal or clinical data

IRB Decision

The Institutional Review Board hereby grants ethical approval and assigns the following code:

CAU-IRB-2025-057

This approval authorizes the researcher to proceed with data collection and to use this code for **academic publication, conference presentation, and institutional reporting**. For verification or further information, please contact: research.ethics@centralasian.uz.

Dr. Sedigheh Shakib Kotamjani
IRB Chairperson

Dr. Sherzodbek Dadaboyev
PhD, IRB Member

Dr. Niladri Maiti, PhD, MDS, M.Sc, BDS
IRB Member

Dr. Sardor Azam,
PhD, IRB Member

Appendix F Cover Letter

Nozima Uralova

264 Milliy Bog street, Mirzo Ulugbek district

Tashkent, Uzbekistan

gs70097@student.upm.edu.my

Dear Editorial Team,

I am pleased to submit the manuscript entitled REFINING QUALITATIVE RESEARCH INSTRUMENTS THROUGH PARTICIPANT INVOLVEMENT: AN INTERVIEW PROTOCOL REFINEMENT PROCESS FOR ELT CURRICULUM RESEARCH IN UZBEKISTAN for consideration for publication in the JIRSEA. This manuscript reports the preparatory phase of a main qualitative research. It mainly focuses on systematic refinement of the interview protocol that is used to explore English Language Teaching curriculum evaluation at higher educational institutions in Uzbekistan. The manuscript is not presenting substantive findings about the ELT curriculum itself, but offers a detailed methodological contribution by describing the multi-stage Interview Protocol Refinement (IPR) process:

1. Matrix development
2. Internal Review Process
3. External Review Process
4. Pilot Study

We believe that this detailed description of the instrument refinement process aligns with JIRSEA's scope, as the journal values rigorous research methods that could serve as a contribution to enhance the field of qualitative research in education. By describing the stages of IPR, we aim to provide practical guidance for researchers who conduct qualitative research in similar contexts as well as novice researchers in the field of education.

We revised the manuscript and removed the primary study's research questions to avoid confusion.

Thank you for considering our submission.

Sincerely,

Nozima Uralova

What It Really Feels Like: Female Academics' Perspectives of Gendered Expectations and Stereotypes in Malaysian Research Universities

ABSTRACT

Authors: Zhang Pu¹, Nurul Afiqah Zulkifly^{1*} & Suriani Ismail¹

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

*Corresponding author:

nurulafiqah@upm.edu.my

Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This qualitative research is an investigation of the role of gender stereotypes in determining the work performance and professional experiences of female academic employees at Malaysian public research universities. The present study is a qualitative research design which takes the exploratory approach, which centers on the in-depth interviews with four female academics working in Malaysian research universities, with a view of generating preliminary information concerning the influence of gendered expectations on academic work experiences. The study, through semi-structured interviews with four female lecturers, reveals that gender bias is manifested in four primary ways: unequal expectations of emotional labour, gendered patterns regarding task distribution, differences in perceived authority, and pressure on women to constantly demonstrate their competence. Based on Role Congruity Theory (RCT), the discussion can indicate that these stereotype-related expectations may serve as impediments to career advancement and appears to reduce job satisfaction due to poor rewarding the contributions of women when acting according to feminine norms and punishing them when they demonstrate agentic behaviours. The findings suggest the need for institutional-level changes that could help to decrease gender inequities and create more equitable academic settings.

Keywords: Gender Stereotypes, Work Performance, Female Academics, Role Congruity Theory, Qualitative Research, Malaysian Higher Education, Emotional Labour, Invisible Labour, Glass Ceiling

1. Introduction

Research shows that although women are represented better than before within higher education, they are still vastly underrepresented in upper levels of academia (European Commission, 2021; Galvin et al., 2024). Perhaps some of this disconnect can be explained by the persistence of gender stereotypes. Gender stereotypes are particularly visible in the academia, where people think of leaders, people in charge, and top researchers only in terms of males. The opposite applies to women, where people think that women are only suited for roles that are supportive, nurturing, and encouraging (Meza-Mejía et al., 2023; Heilman, Caleo, & Manzi, 2024). These stereotypes are not just “theoretical”; they have actual and real impacts to the quality of work and the quality of people’s careers. Administrative and teaching roles that are not valued are attributed to women. In this way, women’s productivity and research visibility are attacked systematically (El-Alayli et al., 2018; Steinþórsdóttir et al., 2021).

In a multiethnic and collectivist country, especially Malaysia, where the expectations surrounding traditional gender roles and stereotypes in work environments are context-sensitive (Hamisan et al., 2023). Malaysia's higher education system is not an exception to the institutional barriers along the lines of the gender. Although women are the majority of personnel in an academic setting, they are still less frequent than men in regard to assuming leadership roles (Ministry of Higher Education Malaysia, 2022; Eaton et al., 2020; Misra et al., 2021). Because of the context in dealing with gender inequalities in the Malaysian academic system, it is imperative that the role of gender stereotypes and how they are an impediment to the effective performance of the employees is examined. Although a number of studies have looked at gender inequities in Malaysian academia (Moorthy et al., 2022; Islam et al., 2023; Devi et al., 2024), few have focused on the women academics and how the stereotypes influence their work performance, professional mobility, and their overall satisfaction with the job. From a practical perspective, the current gap is the lack of sufficient evidence to support the formulation of targeted interventions.

Consequently, university administrators and policymakers lack in deep understanding of these specific stereotypes (i.e. stereotypes related to emotional work, authority and task allocation). But these stereotypes need to be identified to formulate effective and evidence-based diversity and inclusion policies (Lean In. & McKinsey & Company, 2022). Furthermore, from a methodological perspective, previous studies have been using broad surveys or secondary data analysis. This has led to a lack of in-depth qualitative studies that can provide rich and deep descriptions of these female scholars’ narratives (Braun & Clarke, 2021a) which is important in identifying the reasons for these quantitative disparities.

To bridge these gaps, this study uses a qualitative approach. This study aims to identify the gender stereotypes that are commonly faced by female academic staff in Malaysian public research universities and critically examine its impact on their job performance. Accordingly, the study is designed as an exploratory qualitative inquiry that offers preliminary, context-specific insights rather than generalizable conclusions, contributing to theory, practice and methodology towards gender equality in Malaysian higher education. Accordingly, this study was designed to address the following two research questions:

RQ₁: How do female academic staff in Malaysian public research universities experience gender stereotypes in their work environment?

RQ₂: How do these gender stereotypes influence their work performance?

This exploratory, theory-generating study, which is based on four in-depth interviews, aims to surface mechanisms and lived narratives rather than produce representative estimates.

Table 1. The Study's Operational Constructs

Emotional Labour	Gendered expectations that female academics are naturally suited to emotional support, communication, and student care
Gendered Task Distribution	Perceived unequal allocation of teaching, administrative, and service-related tasks based on gender
Authority/Credibility Gap	Explores the phenomenon where female academics perceive their professional authority, expertise, or judgment to be undervalued compared to male colleagues, particularly in meetings and decision-making contexts.
Work Performance	Analyzes performance in domains such as research, teaching, and service, covering key responsibilities and performance metrics across roles.

2. Literature Review

2.1. Work Performance in Universities

In universities, work performance is not just about research output. Although output can and should be measured – publications, amounts of research funding, citations, and so on (Huang et al., 2020) – a more holistic picture must also encompass teaching, administration, and general academic citizenship (Smith et al., 2024). Academics have warned that measuring performance primarily through research output overlooks the many aspects of academic work, such as the hidden role of mentor and servant that maintains university life but is infrequently acknowledged (O'Meara et al., 2020). These many facets of performance are not assessed in a wholly objective way. Institutional biases and cultural influences – male biases in particular – distort assessments of performance, so it is important to examine how organizational culture and implicit assumptions affect them (Heilman et al., 2024).

2.1.1. The Gendered Dimensions of Academic Performance

First, academic work comprises multiple dimensions (research, teaching, service/administration, mentoring, external involvement/influence, etc.), which are not given equal importance in organizational practice (Norander & Zenk, 2023). Recent research emphasizes that institutionalized performance evaluations often prioritize quantifiable research outputs (such as the number of papers, funding, and citations), while significantly underestimating the importance of teaching quality, service, community engagement, and emotional labor (Reid, 2021). This leads to the so-called "invisible labour" problem—that is, a large amount of work crucial to the operation of universities but not fully recognized by the promotion and evaluation system (such as student mentoring, committee work, and unpaid mentoring roles) is often undertaken by women, and this unequal distribution erodes women's research time and promotion opportunities.

Secondly, the study found that teaching and service work are more time-consuming for female scholars. Women are more often expected or requested to undertake pastoral care, student counselling, and departmental/faculty administrative tasks. While these tasks are crucial to student experience and institutional operation, they are typically not included in research-oriented KPIs, thus putting them at a disadvantage in promotion assessments (Sümer, 2023). This creates a structural path from gendered task allocation to limited research output, and then to lower evaluation scores in a system where publications are valued above all else.

Furthermore, student evaluations of teaching (SETs) have been repeatedly shown to be gender-biased in global studies: student ratings of female teachers are often influenced by the teacher's gender, expected role, and classroom behavior (e.g., additional expectations of "warmth/support"). These ratings are frequently incorporated into promotion and teaching performance evaluations, further amplifying institutional disadvantages against women (Khokhlova, Lamba, & Kishore, 2023; Aragón

et al., 2023). Therefore, the teaching dimension is not value-neutral; it is gendered through culture and evaluation mechanisms.

Finally, interdisciplinary and collaborative research, as well as service-oriented research (such as work aimed at educational improvement and community engagement), are often underestimated or fail to translate into traditional academic reputation and quantifiable outcomes. This is more common among female scholars because their research and teaching workloads are more dispersed, further disadvantaging them structurally in terms of quantifiable research metrics (Reid, 2021). Overall, viewing "performance" as gender-neutral masks the profound influence of institutional and cultural factors on job allocation and evaluation.

2.2. Work Performance of Women in Universities in Malaysia

These are also evident in the larger education in Malaysia and may even be more striking due to cultural considerations. As it has been found out, in Malaysia, female lecturers tend to pay more attention regarding the involvement in teaching and service-related tasks than male colleagues in Malaysia (Munusamy, Ramasamy, and Sukir, 2024). It was also discovered that stereotypes according to which men inherently have authority and can better occupy leadership roles exist among the academic staff (Islam et al., 2023). These attitudes may weaken the position of women and reduce valuation of their professional input (Othman et al., 2023). Though these disparities are well established, not much research has dwelled to investigate how these expectations are experienced in day to day academic or how the disparities frustrate the performance and satisfaction of the work they do.

Regarding the issue of higher education in Asian countries, quantitative and qualitative studies indicate similar structural issues: female lecturers are more often given or volunteer to teach more often, and support students and do administrative work in their departments. Although such actions might enhance the quality of teaching and student satisfaction in the short run, they undermine time spent on research and writing work in women during the long run, thereby influencing the production of research and the rate of promotion (Nursyamsi, 2024; Nurumal et al., 2023). In addition, local research indicates that the cultural norms of South Asia and the Southeast Asian people (with the promotion of humility, respect, and hierarchy) increase the demands on women; therefore, becoming micro-excluded or less visible during interactions in classrooms, departmental meetings, and research partnerships (Islam et al., 2023). Such cultural-institutional interaction intensifies the international problem of gendered performance in the situation in Malaysia.

As in most university performance evaluation systems around the world, the Malaysian academic institutions are extremely skewed in their approach toward measurable research results (including publications and granting funding) (Sarjidan & Kasim, 2023). Nevertheless, the academic load is unevenly distributed among female researchers, with a greater proportion having to contribute to housework at school like teaching, serving students, and administrative work (Islam et al., 2023). The withdrawal of women or non-participation in the labor market by Malaysian women is concretely demonstrated in a report by the World Bank (2021): it points out that most women withdrawal or do not enter the workforce because of the so-called domestic work, which involves childcare and elder care. This causes a single-peak trend of Malaysian female labor participation from the working force, which has a large proportion of women dropping out at the age of about 30, the reproduction age, and never coming back.

At the same time, women also encounter barriers to accessing international partnerships and large-scale funding applications due to the influence of informal networks (i.e. mentorship relations and the academic circles) where the resource distribution depends, and women are not involved at all or are poorly represented in these informal contacts (Bellotti, 2022; Callander, Johnson & Grandis, 2025). Thus, the performance of women at work is not taken seriously or is misunderstood in institutionalized ways of evaluation. Lastly, it can be mentioned that since Malaysian society is multi-ethnic (Malay, Chinese, and Indian), female academicians experience varying family expectations and community

standards, and therefore, the so-called work-family balance implies more challenges and coping strategies of the ethnic groups.

2.3. Role Congruity Theory (RCT)

Role Congruity Theory (RCT) (Eagly & Karau, 2002) offers a compelling theoretical framework for explaining the persistence of gender bias in academia. According to RCT, people have shared beliefs or gender stereotypes about the typical characteristics associated with women (e.g., communal traits such as warmth, nurturance), and men (e.g., agentic traits such as assertiveness, dominance). Prejudice occurs when there is perceived incongruity between gender stereotypes and the requirements of a social role.

In the present study, being a senior academic and a leader is often assumed to be agentic. When women, who are stereotypically communal, hold these positions, they encounter a double standard and a penalty for deviation from expected norms. This perceived incongruity between the female gender role and the leader role results in two types of prejudice (Heilman, Caleo & Manzi, 2024): First, women violate expected norms and are judged as less fit for the role; that is, women are evaluated as less capable of holding high-stakes research leadership roles and more authoritative positions because of the perceived incongruity between the female gender role and the leader role. Second, when women do exhibit agentic behaviour to compensate for this negative evaluation, they are evaluated negatively for that behaviour, evaluated more harshly than men for doing the same assertive or decisive things, and labelled as "aggressive" or "difficult".

Hence, RCT explains not just why women face initial barriers to entry into leadership positions in academia, but also why, once there, they continue to face challenges in being taken seriously and in having their performance evaluated fairly. This study uses RCT as a theoretical framework to interpret and analyse the reported experiences of academic women in Malaysia, focusing on how the perceived incongruity between the female gender role and the academic role leads to evaluative disadvantages, and how this affects their work performance.

Building upon the theoretical foundation of RCT, this study constructs a qualitatively oriented conceptual research framework to organize thematic focus during interviews and subsequent coding frames. This framework combines the RCT perspective of "role incongruity" with organizational-level "gendered organizational practices" and individual-level "experience/coping strategies," forming a multi-layered analytical perspective.

2.3.1. From Theory to Analytical Mechanisms: What the Interviews Will Probe

The biases detailed by the authors mentioned above and the outcomes predicted by Role Congruity Theory (Eagly & Karau, 2002) intertwine and present challenges particular to the field of academic work. This study will guide coding and interview development through the biases identified by the authors. Firstly, according to RCT, which proposes two forms of prejudice toward female leaders, these can be further understood as descriptive biases and prescriptive biases. Descriptive biases hold that women are viewed as lacking the so-called 'agentic' qualities that are required of leaders, while prescriptive biases reflect that women are penalized for being assertive. This tension will be the focus of research to be conducted through these interviews. This helps to understand whether respondents believe that women within Malaysian academia suffer from more descriptive exclusion (the perception that women are less suitable for holding a leadership role) or whether they believe that women suffer from more prescriptive penalties (the perception that women are punished for being assertive and taking on leadership roles).

A more recent mechanism that the literature identifies is the invisible allocation of work which, in this instance, refer to the more pronounced allocation of teaching and service work to women (Meza-Mejía, Villarreal-García & Ortega- Barba, 2023; Sümer, 2023), which reduces women's research time and

visibility. This, along with literature on RCT, suggests that this allocation may in fact be a by-product of communal role expectations and biases. These interviews will determine whether respondents believe that the allocation of work is done in a voluntary manner (i.e. through some self-fulfilling prophetic role expectation) or in a manner that is prescribed by the structure within which they operate.

Women are often faced with balancing individual strategies to cope, such as working harder and promoting oneself, and imposing structural change, for example through equitable distribution of work policies. While RCT suggests that individual coping may be self-destructive because of prescriptive penalties, it is often the case that the explanation for backlash is work that is not traditionally done, which in this case is the case for women. These interviews will examine the strategies that women undertake, and determine whether women perceive, or more accurately will themselves suffer from backlash as a result of their strategies. The aforementioned mechanisms served as a basis for the construction of the interview questions (i.e. the allocation of work, the perception of fairness in the evaluation of work done, and the experience of backlash), as well as the coding framework.

2.4. The influence of Gender Inequality Structures on Women's Academic Work

2.4.1. Hidden Labor and Invisible Contributions

A recent empirical study of universities in several different countries demonstrates that male and female faculty members submit the same average amount of weekly hours, yet female faculty members spend more time teaching and serving the university, and men spend more time at the research. Even when this gender difference is controlled by the family and workload, it is still significant (Allen et al., 2023). This conclusion confirms the invisible labor according to which, women do much teaching/service work which cannot be estimated, and they are seldom promoted or evaluated. However, more easily incorporated in the notion of performance, by quantifiable measures, like research output (publication), female scholars are structurally disadvantaged as part of institutional production (like in terms of promotion and research performance evaluation). Moreover, considering external factors like parental leave, maternity leave, and family issues this difference in genders (and, more specifically, the difference in research productivity) becomes much narrower. It implies that one of the gaps is greater non-professional expectations of women (including childbirth and caregiving) and the associated time out of work (Aksnes, Nygaard and Reiling, 2025).

2.4.2. Glass Ceiling

Glass ceiling is an institutionalized and hidden barrier to women getting to the highest or top position in an organization, or getting to a position of power cognized in the literature of management that has over thirty years been discussed (Taparia and Lenka, 2022). Despite the overall increase in the number of women on the general academic faculty, the percentage of women who take up senior positions (associate professor, full professor, senior leadership) is still very meagre. In Europe, an example given is a recent report on the academic systems of the OECD countries: it shows that many of them yet cannot balance gender in the senior level. Women tend to be disadvantaged when it comes to promotions, a chance to receive a research grant, an opportunity to work internationally, the distribution of resources, or involvement in an editorial review or editing panel (OECD, 2024).

To take one case as an illustration, a longitudinal study conducted in Italian universities (2025) demonstrated that despite considering individual productivity and differences in fields, women continued to be greatly disadvantaged during the post-doc to assistant professor transition which also pointed to institutional/structural factors being central in the gender gap (Gaiaschi, 2025).

2.4.3. Work Stress, Mental Health and Job Satisfaction

As institutional and cultural imbalance is historically cumulative, female teachers/scholars have more chances to be held in greater danger of stress, burnout, and dressing. According to a 2025 study, mid-level academic professions show that women experience less job satisfaction and increased stress and burnout rates more often than men because certain structural factors impact the employees, namely overtime, unpaid service, and hidden labor (Röbken et al., 2025).

Accordingly, the studies on the circumstances of the COVID-19 lockdown indicate that female academicians feel contradictions between the demands of work on the family, house cleaning, and providing emotional support to students and their career growth (research, writing) which has a disastrous effect on their research work and employment opportunities (Bam, Walters & Jansen, 2024).

Moreover, in the societies that uphold gender role beliefs (like the Asian societies), women are expected to have a larger portion of the duties in the family (childcare, eldercare), and the household chores. This is superimposed on professional work (teaching, mentoring, administration), which increases work-family conflict and influences mental health and career retention. Such combination of institutional, cultural and personal forces simplifies the situation whereby most of young female academically promising scholars end up quitting academia.

2.4.4. Inequality in Social Capital, Academic Resources, And Visibility

Beyond time and institutional constraints, female scholars are also at a disadvantage in accessing research collaborations, international partnerships, editing/reviewing opportunities, and academic networks. Recent large-scale studies in communication have found that even with a growing number of publications by women scholars, their rates of citations, sole-authored publication, and publication in top-ranked journals remain lower than men's even when they control team composition and research field. These disparities are primarily reflected in research visibility and academic recognition (Oliveira & Huang, 2025).

Meanwhile, research has found that female scholars are also more conservative in self-promotion on social media/public communication; they are less likely to actively promote their research findings (e.g., through Twitter), and even with the same or greater effort, they may receive less reward due to gender bias (e.g., negative evaluations of assertiveness) (Peng et al., 2025). These factors mean that relying solely on publication count or grants as performance/reputation/promotion criteria systematically underestimates the contributions of female scholars especially those whose research is more collaborative, interdisciplinary, or service- or community-oriented.

2.4.5 Inequality in Gender Role Perceptions, Institutional Policies and Structures

A recent review points out that inequality in women's leadership in higher education is not merely a matter of individual ability differences or choice, but rather the result of the combined effects of institutional, sociocultural, and structural factors (Meza-Mejia, Villarreal-García & Ortega-Barba, 2023). These structures include: a promotion system that prioritizes outcomes and lacks recognition of the value of teaching, service and administrative work; implicit preferences in the review, funding and promotion process for social capital, networks, collaboration, and leadership potential; insufficient institutional support for childcare or parenting responsibilities; and culturally stereotypical expectations of gender roles. Therefore, from the perspective of cultural-institutional-structural interactions, even if women and men are formally equal in terms of titles, positions, and qualifications, their "available resources (such as time, energy, and social capital), visibility, and career trajectory" may be systematically eroded in the long term.

2.4.6 Research Rationale and Gap Statement

Based on the latest research, we can see that even though the proportion of women entering university faculty, postdoctoral, and assistant professor positions has increased in recent years, they remain systematically disadvantaged in multiple dimensions, including resource allocation, implicit labor, academic visibility, psychological burden and risk of attrition, and promotion and leadership opportunities.

However, most current research on these issues is based on quantitative methods (such as surveys, bibliometric data, and time allocation statistics) and is concentrated in Western countries (Europe and North America) or a few countries with publicly available data—research on the institutional and cultural interactions in Asia, especially in multi-ethnic and multi-racial societies (such as Malaysia), is still relatively limited. As the structural review points out, gender inequality manifests in different forms in different countries, cultural, or institutional contexts, and the impact of institutional reforms (such as institutional support, childcare policies, work flexibility, and promotion mechanisms) on mitigating inequality has not been fully explored.

Furthermore, even when research focuses on work-family conflict, childcare responsibilities, attrition rates, and burnout, it primarily focuses on basic education or teaching-oriented universities and may not be applicable to research-intensive universities. For example, research on burnout among female teachers is mostly conducted in teaching-oriented universities or the basic education system (Wei, Subramaniam & Wang, 2025). Therefore, exploring the implicit labor, time allocation, and resource access of female academic staff in Malaysian public research universities through qualitative interviews and a multi-layered framework (i.e., institutional, cultural, and personal experience) fills a gap in the current literature.

3. Methodology

3.1. Research Design

To answer the research questions, exploratory qualitative research method was adopted to get insights on the common gender stereotypes and their impacts on female academic staff in Malaysian public universities. The research design involved collecting and analyzing data on female academic staff in public universities with semi-structured interview and document.

Interviews can yield full and rich accounts of how people see the world, what sense they make of it, and what concerns they bring to their lives (Demirci, 2024). Different from a highly-structured interview that rigidly adhering to predetermined questions, which may not allow the researchers to access participants' perspectives and understandings of the topic, semi-structured interview, which is guided by a list of questions (Ruslin et al., 2022), was used to understand the range of perceptions of gender stereotypes by female academic staff in Malaysian public universities. Therefore, the present study utilized semi-structured interview as its primary data collection method.

A pilot interview was conducted online with a female university lecturer to ensure that questions are worded appropriately and are understood and interpreted as intended (Lim, 2025). As proposed by Morgan (2022), data from document can elicit meaning, gain understanding and advance new categories. Stability is another advantage of using documentary material. Unlike interviewing and observation, the presence of the investigator does not alter what is being studied. Documentary data are “objective” sources of data compared to other forms. In this study, documents were used to do documentary contextualisation of the data gleaned from the interviews. The whole research flow is illustrated in Figure 1.

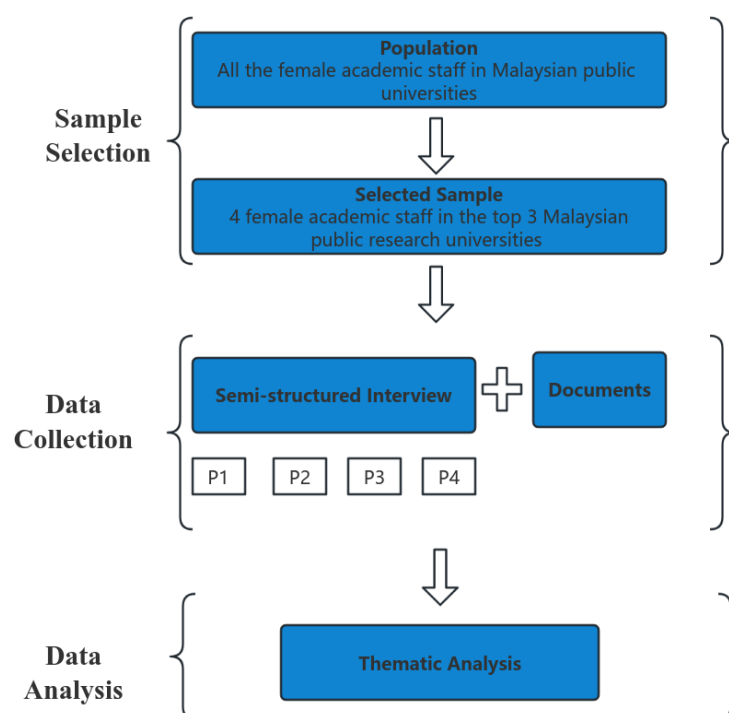


Figure 1. Data Collection and Data Analysis Flow

3.2. Participant Information

This study was conducted in three of the top public research universities in Malaysia through purposive sampling. The sampling frame comprised public, research-intensive institutions representing diverse disciplinary profiles. Potential participants were identified based on predefined inclusion criteria, including gender (female), current academic appointment, and a minimum of two years of academic experience.

Recruitment invitations were sent directly by the researchers to eligible participants identified through publicly available institutional websites and professional networks. The invitations outlined the study purpose, voluntary nature of participation, and confidentiality assurances. To collect the data, the period was August 2025 to October 2025. Although they could invite eligible academics again and again during the time (23 direct invitations were sent to qualified academics), the number of eligible academics that were willing and available to participate was rather low.

Finally, four respondents accepted and undertook to fill the in-depth interviews. Because of the depth-oriented and exploration of the study, the final sample was regarded as a rich set of information and not a representative set. Such a recruitment result is well aligned with qualitative pilot studies which are concerned with the depth of analysis rather than the sample size. Table 1 represents the criterion of the sampling, and Figure 2, Table 2 illustrates the entire procedure of the recruitment process and participant flow, respectively:

Table 2. Criterion for Sampling

Criteria	Specifics
Gender Requirement	Must be female academic staff
Employment Status	Must be currently employed as academic staff Should hold a formal academic position

Institution Type	Must be employed at a Malaysian public research university Not private universities or other higher education institutions
Employment Duration	Should have worked in the academic field for 2 years at least
Academic Disciplines	From various academic disciplines
Career Levels	From different academic positions

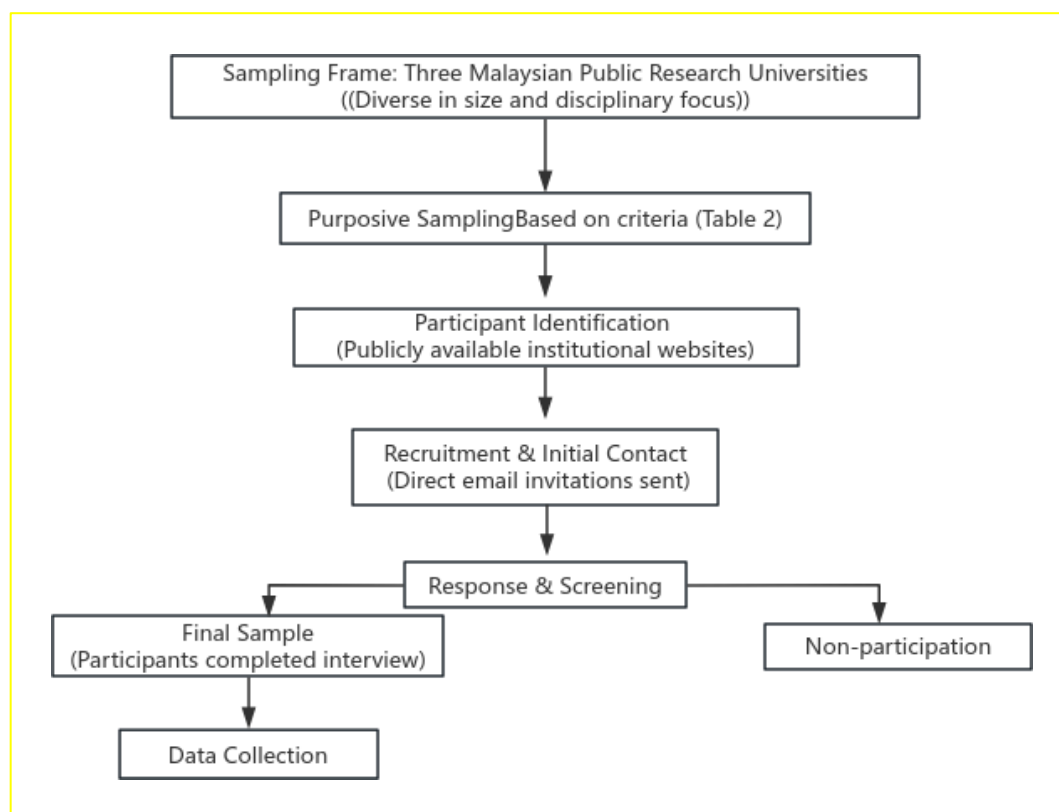


Figure 2. Recruitment and Participant Flow

Table 3: Demographic and Data Collection

Serial Number	Position	Cultural Background	Experience	Discipline	Method	Camera	Duration
P1	Senior lecturer	Malaysian	6 years	STEM	Zoom	On	46 mins
P2	Senior lecturer	Indian	20+years	Non-STEM	Zoom	On	66 mins
P3	Senior lecturer	Malaysian	16 years	Non-STEM	Zoom	Off	52 mins
P4	Lecturer	Malaysian	5 years	Non-STEM	Google meet	On	55 mins

3.3. Data Collection

The online interviews were done individually through Zoom or Google meet. All the interviews were recorded and transcribed word-to-word to analyze them. Table 2 demonstrated the time and the number of words spent on every interview. All the participants were provided with written consent forms. The gathering of documents was done following the technique suggested by Guest, Namey and Mitchell

(2013). To begin with, we set inclusion criteria: we determined what was included in the scope, what type of documents, and what was included within this period of time. Second, we gathered materials in the form of documents in the open or organizational sources.

Similarly, to the exploratory character of the research in question, data collection did not target exhaustive thematic saturation but preliminary, in-depth coverage of the phenomenon of gender stereotypes experienced by female academics working in Malaysian public research universities. The sampled population was purposely chosen and belonged to a rather homogenous group and presented intensive, reflective information concerning their careers. In the course of analysis, common trends associated with gendered expectations, emotional labour, perceptions of authority and distribution of tasks were discovered to be recurrent in interviews. However, it was by the fourth interview that no significantly new themes were identified in the context of the research question that the data was seen to be adequate in the exploratory and theory-making purpose of this study.

3.4. Data Analysis

As the main approach to the analysis of the collected data, thematic analysis was used, it is the process of identifying a pattern, which triggered the emergence of the themes and placed them into categories to be analysed further. This included attentive and dedicated reading and re-reading of data, systematic coding and category building (Braun and Clarke, 2021a & Braun and Clarke, 2021b). Besides interview data were a qualitative document analysis (QDA) that was performed to aid documentary contextualisation. The qualitative document analysis was constrained by three factors used to choose the documents to be included in the analysis. To begin with, they had to be authoritative and publicly accessible sources that were created by established governmental or international organisations. Secondly, the documents were required to contain empirical or policy-related information of gender, academic work or labour participation that would have been contextually relevant to the study focus on academic females. Third, only the currently existing or still ongoing documents (2020-2025) were included, and the information should have matched what the participants were reporting about. On these criteria, 4 documents were selected (see Appendix C). These documents were coded using the deductive approach to coding proposed by Guest et al. (2013) and Morgan (2022) to investigate the perceived themes in the data of the interviews. This research allowed to make a direct comparison between formal institutional expectation and self-reported lived experiences of the participants.

To test the analysis, intercoder reliability guidelines were used as a sense of verification. The initial coding procedures were formed via online meetings by the researchers to create a complete codebook. We then analysed transcripts and documents separately discussing any gaps in coding to bring to a consensus. This incorporation of document insights in the meaning making of the findings increased the depth and credibility of the study in general. Member-checking was done to guarantee the worthiness of the results. Following the preliminary analysis, the subjects were given reports about the preliminary themes and essential results through email. The thematic descriptions in the category "gendered task distribution and inequalities in workload" are slightly altered as a result of the feedback provided by two respondents to portray the subtlety of administrative burden.

The research team agrees that we are all bound to have influenced the process of interpretation through our backgrounds. As female researchers (First author), and female scholars (Second and third author) operating in the Malaysian higher education field, we enjoyed an insider view, with the advantage of rapport and trust with the respondents. Nonetheless, this common experience also gave rise to a danger of a pre-existing bias, as individual dissatisfaction with workload might conceal the data. To overcome this, we kept a reflexive journal during the coding procedure and had peer debriefing sessions on a regular basis. These debates compelled us to question our original expectations and made the findings firmly based on the voice of the participants instead of our identity.

3.4.1. Saturation Assessment and Stopping Rationale

Given the exploratory and theory-generating aim of the study, coupled with the small sample size ($n = 4$), it is important to explain how saturation was measured and why this sample is regarded as sufficient. Saturation was measured according to the code frequency counts method described by Hennink and Kaiser (2022).

After each interview, the research team independently analyzed the transcript and then convened to discuss their codes. The first interview produced 12 codes. The second interview produced 4 new codes. The third interview produced only 1 new code. The fourth interview produced no new codes apart from those already described in the first three interviews. This pattern of no new codes in the last interview—was our stopping rule (Hennink & Kaiser, 2022). The code accumulation matrix (Appendix E) illustrates the progress of this process.

Code saturation can be achieved through a lesser number of interviews when the study population is homogeneous, and the research aim is specific (Hennink & Kaiser, 2022). While our sample of four participants is small, it was purposively drawn from a homogeneous population: female academics in public research universities in Malaysia, with a minimum of two years of professional experience. The participants had similar institutional settings, career levels (three senior lecturers and one lecturer), and professional challenges. This homogeneity, in conjunction with the depth-oriented nature of the interviews (each lasting 45 - 66 minutes), enabled us to collect a broad and comprehensive range of data with a smaller number of interviews.

The focus of this study is neither to make statistical generalisability claims nor to attempt to verify existing theories. This study looks to create a hypotheses and describe the workings of gender stereotypes in this context. A focused, small sample qualitative study is methodologically fitting for this purpose. While larger scale studies are required to confirm and broaden these findings, for the theory generating aims of this exploratory study, the sample size is rational.

3.4.2. Coding Process and Consensus

The coding process used thematic analysis as explained by Braun & Clarke (2021b), which involves cycles of coding and theme creation. The first step for the researchers was to familiarize themselves with the data and establish preliminary codes in the interview transcripts. These initial codes went through cycles of review and refinement which led to the creation of larger category codes that eventually formed the final themes.

In order to increase the consistency and trustworthiness of the coding, two researchers coded the same transcripts independently and then reconciled their coding. The research team held three consensus meetings that lasted around 60 to 100 minutes. During these meetings, the two coding team members shared the results of their independent coding, and where there were discrepancies (for example, where one coder assigned a code and the other coder did not), they discussed the case and resolved it by consensus. When there was a stalemate, a third researcher was brought in to resolve the issue (there were two instances of coding where this happened). The coding consensus meetings also provided the basis for the iterative development of the codebook and the refinement of the codes in the manual, which was done in subsequent meetings. Appendix B includes a sample of the codebook that was used.

3.5. Variation and Moderators

When going beyond description and examining possible reasons for variability between participants, the four female academics were allocated to the primary analytical dimensions from the literature. This can be found in Table 4.

Several of the tentative patterns from the cross-case comparison are as follows. Regarding the discipline, the only STEM representative (P1) did not cite greater time pressure to produce research outputs compared to her non-STEM colleagues. Rather, she benchmarked herself against her male colleagues,

stating she felt more pressure than they did. This suggests that a gender comparison within the same discipline may be a more relevant perspective to consider than the discipline as a whole to understand stereotype impacts.

In addition, the position in the career ladder seemed important. The early-career participant (P4, lecturer) said she found it much harder to refuse teaching and service tasks than her more senior colleagues. This suggests that those with less institutional seniority feel the effects of invisible labour more. Unlike her, it appears that senior lecturers (P1, P2, P3) have less of such concerns, which could be a result of the greater autonomy or bargaining power they have gained over time.

P2 and P4 referred to caregiving responsibilities, but the female academics primarily structured their explanation as a comparison between the genders. They pointed out that women, in contrast to their male colleagues, have additional unpaid caregiving responsibilities at home. Neither participant provided many details on work-family conflict. This signifies that the claimed injustice is more prominent in the gendered allocation of unpaid domestic work than in the conflict between work and family roles.

Lastly, concerning P2's ethnicity, the only Indian participant, she stated that she did not encounter any incidents of marginalisation or exclusion attributable to her ethnicity. It could be the case that ethnicity did not constitute a significant aspect of her experience or that the professional milieu of the university is sufficiently inclusive. Given the limited sampling, no definitive conclusions can be made.

These considerations are very preliminary given the sample size and the fact that not all of the participants addressed every single dimension. Instead of generalising, this cross-case mapping is intended to formulate hypotheses and indicate areas for future research on a larger scale.

Table 4. Participant Mapping on Key Dimensions

Serial Number	Career Stage	Ethnicity	Discipline	Caregiving Status
P1	Senior lecturer	Malaysian	STEM	Not mentioned
P2	Senior lecturer	Indian	Non-STEM	Yes
P3	Senior lecturer	Malaysian	Non-STEM	Not mentioned
P4	Lecturer	Malaysian	Non-STEM	Yes

3.6. Validity and Reliability of Data

To ensure methodological rigour, several strategies to enhance validation of the study were adopted (Bryant & Ballantyne, 2017) as summarised in Table 3. Data were subjected to documentary contextualisation by obtaining evidence from the semi-structured interviews with the participants, to see if this evidence was supported by the documentary sources. Descriptive details of the research setting and research process were provided to enhance external validity. Reliability was maintained by investigator using secondary data for contextual grounding. The multiple researchers conducted an independent coding and analysis of the documentary sources and interviews.

Furthermore, researchers' subjectivity was weakened via reflexive actions as we attempted to monitor our own assumptions and biases during the research process. Finally, all the above steps led to the scientific validity and credibility of our study.

Reflexive vignette 1: We were all female academics in Malaysian public universities, which increased the risk of over-identification with participants. We purposely looked out for disconfirming evidence in transcripts and sometimes questioned whether each code was reflecting participants' voices or our expectations.

Reflexive vignette 2: One of the researchers coded more emotional labour codes than the other did during the coding process. To minimise this bias, we returned two preliminary theme summaries to participants for member-checking, and these participants affirmed our interpretations. We also had a peer debriefing with a colleague outside the team who criticised us to make us justify every code with evidence in the transcripts.

Table 5. Trustworthiness Framework

Criterion	Strategy employed
Construct Validity	-Documentary Contextualisation
	-Member check
External Validity	Thick description
Reliability	Using Secondary Data for Contextual Grounding
Objectivity	Practice reflectivity

3.7. Ethical Approval

Ethical approval for this study was obtained from the Ethics Committee for Research Involving Human Subjects, Universiti Putra Malaysia (Jawatankuasa Etika Universiti Penyelidikan Melibatkan Manusia, JKEUPM). The study was approved under reference number JKEUPM-2025-015, with approval validity from 27 August 2025 to 26 August 2026. Prior to participation, all participants were provided with detailed information about the study's purpose, procedures, and their rights as research participants. Written informed consent was obtained from all participants before data collection commenced. Participation was voluntary, and participants were informed of their right to withdraw at any stage without penalty. To ensure confidentiality, all interview data were anonymised through the use of pseudonyms, and any identifying information was removed during transcription. Audio files and transcripts are stored in a password-protected shared cloud folder with access restricted to the research team. They will be retained for five years following the completion of the study and will be permanently deleted thereafter. Anonymised data are available upon reasonable request to the corresponding author, subject to institutional ethical approval.

4. Findings and Discussion

As an exploratory study, the findings should be interpreted as theory-generating rather than theory-confirming. With a thorough thematic analysis of collected interview data, four different but interrelated themes were identified as summarized in Table 4, providing an in-depth understanding of prevalent gender stereotypes, how gender stereotypes impact their professional experience, career advancement and job satisfaction in educational setting. This section presents a detailed discussion of the research findings, supported by relevant quotes from the interviews.

Table 6. Summary of Findings

Theme	Key Manifestations	Perceived Impact
Emotional Labor Expectations	-Assumption that women naturally excel in communication and emotional support (P1).	-Increased workload and work pressure. -Emotional burnout and mental strain.

and The Invisible Burden	-Routine assignment to student counseling and pastoral care duties (P3). -Male colleagues are assigned more technical roles (P3). -Students expect more empathy from female teachers (P4).	-Reduced job satisfaction and overall well-being. -An invisible, unappreciated burden is placed on female educators.
Gendered Task Distribution and Inequalities in Workload	-Women are expected to take on administrative duties (P3). -Men are channeled towards senior leadership roles (P3). -STEM fields are perceived as male domains (P4). -Women are assigned event planning, paperwork, and committee work. (P3)	-Limits women's capability to engage in academic and research activities. -Creates unequal workload and opportunity. -Perpetuates gender disparities in leadership and STEM fields. -Results in gender-based bias in promotions .
Authority Perception Disparities and The Credibility Gap	-Male colleagues' ideas are prioritized in meetings (P2). -Students expect male teachers to be more assertive and disciplined (P3). -Female authority is frequently questioned (P2).	-Undermines confidence and effectiveness in leadership positions. -Negatively affects female educators' confidence. -Creates a challenge in winning the same respect as male colleagues.
Constant Need to Prove Competence and Career Advancement Barriers	-Capabilities are frequently questioned, leading to a need to overprove oneself (P1). -Contributions are undervalued or overshadowed by male counterparts (P2). -Being overlooked for projects in favor of male colleagues (P4).	-Slower career advancement due to double standards and lack of recognition. -Diminished job satisfaction and weakened sense of fulfillment . -Creates extra pressure and exhaustion. -Contributes to the underrepresentation of women in senior roles.

4.1. RQ1: What are the gender stereotypes experienced by female academic staff in Malaysian public research universities?

This section presents the gender stereotypes experienced by female academic staff at a research-intensive public university in Malaysia. Through interview analysis, three main themes emerge, revealing how organizations, students, and colleagues influence women's daily work and role positioning through gendered expectations.

4.1.1. Emotional Labor Expectations and The Invisible Burden

Emotional labour emerged as the most prominent theme in our findings. Three of four participants (P1, P3 & P4) mentioned that they were often expected to control their emotions and provide interpersonal support in workplaces.

"You know, female teachers often face assumptions that they should naturally excel in communication skills and emotional support, while males are viewed as better at discipline and leadership." (Participant 1)

Two of respondents (P3 & P4) mentioned that women educators were usually expected to play an emotionally demanding role such as counselling students or pastoral care. In addition, these extra works usually increased their workload but had no extra rewards.

"Well, I have witnessed my female colleagues often being asked to handle student counseling, while male colleagues are assigned more technical roles." (Participant 3)

“Students may expect female teachers to be more empathetic and understanding, while perceiving male teachers as figures for discipline and structure. For instance, girls might seek emotional support from female teachers, which can be stressful as it adds pressure.”

(Participant 4)

This theme will align well with a certain literature that suggests that emotional labour is an invisible weight on women in academia which is within the context of this study. The narratives by the participants indicate that, female scholars were always supposed to assume pastoral roles, and mentor the pupils and handle interpersonal problems among other students. This type of communal behaviour though aligned with gendered expectations is not promoted in promotion structures that still place emphasis on research productivity. This is a reflection of the more recent studies which have indicated that emotional labour is overly perceived as female, and it is not sufficiently rewarded or acknowledged (Reid, 2021). The resulting trend of these demands may overwhelm female scholars, which may be associated with to exhaustion (Kariou et al., 2021) and could limits their performance in research.

4.1.2. Gendered Task Distribution and Inequalities in Workload

Gender differences in the allocation of work were evident among the interviewed group. Participants (P3 & P4) repeatedly mentioned that the organization would follow the traditional gender division of labour.

“Emmm, gender stereotypes manifest in my working environment through expectations that female colleagues take on administrative duties, while male colleagues are expected to pursue senior leadership roles.” (Participant 3)

As mentioned by participants (P1 & P4), girls usually got less encouragement or support from the institution to apply for leadership or STEM positions, which may have connection with the phenomenon of gender imbalance.

“Yes, certain subjects are often seen as more suitable for one gender. For instance, STEM fields are typically associated with males, while arts and humanities are considered more suitable for females. (Pausing) This bias likely stems from societal expectations and traditional roles that shape perceptions of competence.” (Participant 4)

In the educational context, as reported by most participants (P1, P2 & P3), some female lecturers were more likely to be assigned to organization events, academic documents, student affairs and committee works while male educators were more likely to handle strategic planning, technical responsibilities and resource-related decisions. For the women in this cohort, this pattern indicates how gender-based role expectations further limit women’s participation in research and scholarly development.

“Yes, I believe there is a gender-based bias in promotions and leadership opportunities. (Pausing) Women are often overlooked for these roles due to prevailing stereotypes about their capabilities and commitments, which unfairly skews opportunities in favor of their male counterparts.” (Participant 3)

Such a tendency in the distribution of tasks according to gender supports the constant unequal results in performance. As our interviews indicate, women were much more often than men assigned to administrative, service, and event-related duties, as men were more likely to be encouraged or oriented towards administrative/leader or research-intensive jobs. These categories seem to indicative of the persistent stereotypes of women and men as being respectively supportive and agentic, respectively. These trends continue to contribute to leadership inequalities in academia, as they have been emphasized in recent scholarships (Misra et al., 2021). As a result, women do not have as many chances

to increase their research visibility and become more prominent in their research and make an academic difference, as well as, slack the career growth based on merit.

4.1.3. Authority Perception Disparities and The Credibility Gap

Among the participants interviewed, gender-based perceptions of authority in the educational environment differ significantly. Participants (P2 & P3) reported being overlooked for leadership positions because they are believed to be less authoritative and less decision-making role models.

"Well, male colleagues are perceived as more authoritative. Once, during a departmental meeting, a male colleague's ideas were prioritized over mine despite similar merit." (Participant 2)

According to the accounts of four participants, gender stereotypes about authority tend to undermine the authority of women teachers in areas that require assertiveness.

"Yes (nodding), you know, students often expect male teachers to be more assertive, while they anticipate female teachers to be more patient. For example, boys may allow themselves to be less disciplined in female-led classes. " (Participant 3)

In these specific instances, women struggle to win the same respect and recognition as their male colleagues which in turn affects their confidence in their leadership role. Being overlooked or interrupted in the workplace setting tends to have a negative impact on the female educator's confidence.

"(sigh) Colleagues sometimes questioned my authority in discussions, making me more frustrated and also determined to assert myself despite these challenges." (Participant 2)

The impact of role incongruity can also be demonstrated by inequality of power and credibility. In their interactions with students and peers, college leaders were more likely to assign more subject-matter knowledge and intellectual authority to male scholars and undermine the achievements and leadership of female scholars. This goes in line with the fact that evidence still points to leadership qualities still being culturally coded in a masculine way (Manzi, Caleo, and Heilman, 2024). Consequently, female scholars are confronted with discriminatory classroom experiences and lessen of perceived competence which overtime might erode confidence in their career and leadership positions.

4.2. RQ2: How do these gender stereotypes influence their work performance?

This section examines how stereotypes affect the work performance, career paths, and job satisfaction of female academics. Analysis suggests that these effects are multi-layered and cumulative, and are reflected in areas such as promotion opportunities, leadership development, and psychological experience.

4.2.1. Constant Need to Prove Competence and Career Advancement Barriers

Another finding from the interviews was that all four participants often felt they had to constantly prove themselves to gain recognition. Participants (P1, P2, P3 & P4) mentioned that they felt distressed about the emotional and cognitive burden of having to prove themselves to counter gendered assumptions.

"Gender stereotypes have often led to assumptions about my capabilities, impacting my confidence and pushing me to overprove myself in professional spaces (silence). I mean, their presence has influenced my career trajectory, compelling me to navigate obstacles that male counterparts may not face." (Participant 1)

This statement describes the need for women to constantly prove themselves and to shake off doubts about their abilities.

"My contributions were either undervalued or overshadowed by male counterparts. Indeed, these experiences have persisted, occasionally affecting my confidence and professional relationships." (Participant 2)

In addition, P4 also mentioned that their achievements were required to be stricter or different standards.

"I've faced instances where my capabilities were questioned due to my gender, often being overlooked for projects in favor of male colleagues. This has affected my visibility and recognition in my field." (Participant 4)

Because of such double standards, female academics often experience slower career progression which is associated with slower promotion and job satisfaction for female academics. The need to constantly prove themselves also can lower their sense of achievement.

"Gender stereotypes negatively impact my job satisfaction by creating an environment where I have to continuously prove my worth. I mean, this additional pressure undermines the joy of achieving my goals." (Participant 1)

The Ministry of Higher Education (MoHE, 2022) states that in the public universities of Malaysia, over half of the workforce is made of women but only approximately 35% of professors. This current gap indicates the systemic limitations to female progress as well as suggests the very fact that the gender stereotypes remain a restricting factor to equity and diversity in higher education.

Female academics in this study described a recurring prove-it-again dynamic in which demonstrating competence required greater and more sustained effort than that expected of male colleagues for equivalent institutional recognition. They were not given due credit or their actions were not considered when performance was in same calibres. Recent research points out that this two-tier academic assessment and rewarding still persists (Llorens et al., 2021), thus resulting in slow career advancement and an underrepresentation in high positions.

"Many women experience bias not only because of their gender, but also because of their race, sexual orientation, a disability, or other aspects of their identity—and the compounded discrimination can be much greater than the sum of its parts. As a result, these groups of women often experience more microaggressions and face more barriers to advancement." (Lean In., & McKinsey & Company, 2022, p. 20)

While participants in this study did not report compounded discrimination based on ethnicity, the Women in the Workplace 2022 report (Lean In, 2022, p. 20) documents that such dynamics exist elsewhere. The Indian participant in this study (P2) did not report marginalisation based on her ethnicity, which may reflect the small sample size. This highlights the need for future research with larger, more diverse samples to explore whether ethnicity amplifies career advancement barriers in the Malaysian context.

These four themes together provide an indication of how the issue of gender stereotypes may be entrenched in the practices of the studied Malaysian public research universities. According to Role Congruity Theory (Eagly & Karau, 2002) this type of incongruity creates both descriptive and prescriptive forms of bias (how women are viewed and how women are expected to act respectively). The evidence from this exploratory sample suggests that the solution to these structural inequities could involve institutional reforms including the sufficient proper valuation of emotional and service work, the fair distribution of tasks, and glimpsing promotion. Unless these changes take place on a systemic level, the rise in female representation will probably be superficial, and it is likely that the performance and career progression obstacles identified in this study may continue to exist.

4.3. Disconfirming Vignettes

While the four participants' narratives were largely convergent, one minimally conforming vignette emerged that merits attention.

The sole Indian participant (P2) described an episode where she resisted having her ideas dismissed by her male supervisor:

"I remember this one time... I had a disagreement with my supervisor... he's male, by the way... and, you know, usually in situations like that, if you just follow the normal process, my ideas probably wouldn't have been accepted. But I just... I kept insisting. I really believed in what I was proposing, so I didn't back down, even though it was quite uncomfortable. Honestly... the whole process wasn't very pleasant... there was tension, and it felt quite difficult at times... But in the end... I did manage to get the outcome I was aiming for... so I guess it was worth it."
(P2)

This vignette is minimally conforming because, compared with the passive acceptance of authority, which was not passively accepted by other participants, P2 actively resisted and had a positive outcome. However, it also acknowledged the emotional toll ("unpleasant", "tension", "quite difficult"). This suggests that not all women passively internalise stereotypes and that agency does not come for free when facing gendered barriers.

No one reported experiencing supportive leadership or resisting gendered role allocation without any emotional cost. This may reflect the power of gender stereotypes in the studied context.

5. Discussions and Conclusion

The results of this research indicate that gender stereotypes may be very strong, even within the academic culture of the investigated Malaysian publicly gathered research universities, and thus reflect the working reality of the female scholars used in the study on continuing and consequential terms. The four themes identified, which were repeatedly reported in the interviews indicate that gendered norms can be replicated in the daily interactions of organisations and institutional practices. They are not individual problems of individual bias; the statistics write tendencies in the frameworks that could hold women at the positions that are both necessary and underestimated, and make the working environment where work based on the feminine role is accepted and a work based on leadership role is audited.

A serious implication of the patterns, as voiced by the participants (P1, P2, P3 & P4) is that it had a direct effect on women thereby making them unable to fully engage in academic work, which is the fuel of career growth. When female scholars in this generation are overburdened with administrative tasks, student service, and other types of invisible labour, they complain that they lose time, attention and visibility in institutions that allow them to be fruitful in their research and relevant in the leadership front. Simultaneously, respondents pointed at a persistent prove-it-again load, according to which they are interrogated about their competence and authoritarianism, and in order to become as legitimized as their male peers, they must exert extra effort, both emotional and cognitive. All these experiences, based on the stories gathered, seem to destroy job satisfaction, narrow leadership paths, and destroy psychological health.

Despite the fact that the involved universities tend to represent themselves as a meritocratic environment, the results depict how informal rules might influence academic life in gendered means. This evidence based on this study indicates that opinions regarding the allocation of workload, the authority of women could be considered credible to a significant extent by those involved in the case which might have been caused by the deep-seated stereotypes. Such difference between what is written in the policy and what is experienced by the participants suggests a major institutional disconnect within these situations. Failure to challenge these unspoken rules means that universities may unknowingly

reproduce a culture where women are only visible by falling into the role of feminine expectations and a made suspect when they take charge and challenge existing hierarchies.

Notably, these results provide a positive contribution to the cause which argues that gender equity cannot be attained solely at an individual level. Teaching women to be more aggressive, tough or strategic do somehow help alter structural conditions that reproduce inequality. The examples given by the participants are pointing to the fact that the problem is not only a personal experience, but also the institutional mechanisms which seem to legitimize unequal allocation of labour. Should organisational culture reward masculine-coded behaviours and undermine relational or administrative inputs, the results suggest that female academics might still encounter systemic and not personal barriers.

To sum up, on the basis of the findings of this exploratory research, this study finds that in light of the intensity of this research, universities should ideally cease to have only symbolic commitments to gender equality, but introduce reforms to their structures. Findings reveal that the solution to inequality issues seems to cover opening workload allocation systems, incentives of emotional and service labour, fair promotional schemes as well as proactive measures to defy stereotype-driven assumptions on power and ability. Unless institutional measures are taken, in terms of deliberate and sustained action, the numbers of women in the academic might be hard to destabilise the structures that marginalise them. To make female academics flourish in real sense, universities can take a critical look into and challenge the norms, practices and power dynamics which were found in the course of this study.

5.1. Expand the Research Agenda and Next Steps

These exploratory findings enable concrete next steps, which are three in number. First, replicate the study in a larger qualitative sample (20–30 participants) across more faculties to generate broader disciplinary variation. Second, apply a supplementary survey derived from the four themes to establish their prevalence in a broader population (200–300 respondents). Third, use a longitudinal cohort study design to follow women academics for 3–5 years and investigate whether gendered stereotypes can help predict attrition and other career outcomes.

On the implementation side, institution-specific actions include transparent workload tracking, unconscious bias training and targeted mentorship programmes — all of which can be implemented in 0-2 years. System-level policy reforms (e.g., reforming promotion criteria, requiring annual gender equity audits, decoupling leadership from masculine-coded traits) require broader coordination and will likely take 3–5 years or longer.

6. Implications and Recommendations

These recommendations in this following section are presented as hypothesis-generating directions for future research and policy consideration rather than evidence-based prescriptions. The results indicate that to curtail well established gender stereotypes, universities can be able to go beyond the symbolic investment in tangible structural and procedural changes. The policies may need to reward and acknowledge emotional labour and service work, even distribution of teaching and administrative work and impose a clear and criteria-based promotion and recognition measures. Devoid of such systemic modifications, merely filling the academia with more women is, in and of itself, unlikely to demolish the very system that keeps on reinforcing the disadvantage.

Universities can consider to invest on special mentorship and leadership development programmes for female academics which might have chance to give them access to guidance, networks and knowledge institutional basis that can lead them towards research-based and leadership positions. Unconscious bias and gender equity Professional development activities, scheduled regularly, including both academic and administrative personnel, can increase awareness and break the inequitable pattern of hiring, promotion, and workload distribution.

The limitations on transferability are the exploratory design and small, purposively chosen sample but they do give a basis to carry out further studies of bigger magnitude of qualitative study or a mixed-method study. Future research should examine whether a broader sample that includes academics from various disciplines, ranks, and institutional types. A combination of the qualitative and quantitative methods could provide more comprehensive information on the workings of gender stereotypes in diverse situations, as well as on the most effective mechanisms of empowering female scholars and promoting gender equity within institutions.

Finally, these findings suggest university leadership may be able to devote themselves to ensure that gender equity policies are systematically monitored and reviewed not only through the tracking of representation but also through the survey of tendencies in work load assignment, promotion results and recognition of contribution. Universities may consider offer the opportunities needed to achieve desired results through concerted and prolonged institutional action, in which female academics may be able to strive and succeed without facing disproportionate disadvantages. In order to give a strategic roadmap of the institution leadership, table 5 will prioritize the recommendations of the study into short, medium, and long term activities and what the primary owners and key performance indicators (KPIs) of various initiatives will be.

Table 7. Prioritized Policy Table

Priority	Strategic Recommendation	Primary Owner(s)	key performance indicators (KPIs)
Immediate (0–1 Year)	Transparent Workload Tracking: Implement systems to formally record and recognize "invisible" service and emotional labor	University HR; Faculty Deans	Introduction of a digital workload dashboard; 100% logging of pastoral care hours.
Immediate (0–1 Year)	Unconscious Bias Training: Conduct professional development sessions for hiring and promotion committees.	Center for Academic Excellence (CAD); HR	Certification of 90%+ of committee members; pre- and post-training equity awareness scores.
Medium (1–3 Years)	Mentorship & Leadership Programs: Launch targeted development programs providing access to networks for female academics.	Office of Research & Innovation; VC's Office	20% increase in female-led research grant applications; increased female representation in senior roles.
Medium (1–3 Years)	Criteria Reform: Explicitly integrate "service" and "interpersonal support" into formal promotion and reward rubrics.	University Senate; Ministry of Higher Education (MOHE)	Updated official Promotion Handbook reflecting broadened merit criteria.
Long-term (3+ Years)	Systemic Monitoring & Audit: Establish annual gender equity audits tracking workload distribution and promotion outcomes.	Vice-Chancellor (VC); Institutional Quality Assurance	Publication of annual Gender Equity Progress Reports; narrowed gap in gendered workload distribution.
Long-term (3+ Years)	Structural Culture Transformation: Decouple leadership expectations from masculine-coded traits through institutional policy shifts.	Top-level University Leadership	Longitudinal staff surveys showing significant improvement in perceived organizational equity.

References

- Aksnes, D. W., Nygaard, L. P., & Reiling, R. B. (2025). A matter of time? How absence from work affects gender gaps in research productivity. *Higher Education*, 1–16.
- Allen, T. D., Miller, M. H., French, K. A., Kim, E., & Centeno, G. (2023). Faculty time expenditure across research, teaching, and service: Do gender differences persist? *Occupational Health Science*, 7(4), 805–818.
- Aragón, O. R., Pietri, E. S., & Powell, B. A. (2023). Gender bias in teaching evaluations: The causal role of department gender composition. *Proceedings of the National Academy of Sciences*, 120(4), e2118466120.
- Bam, A., Walters, C., & Jansen, J. (2024). Care and academic work in a pandemic lockdown: A study of women academics in South Africa. *Higher Education*, 87(6), 1791–1807.
- Bellotti, E., Czerniawska, D., Everett, M. G., & Guadalupi, L. (2022). Gender inequalities in research funding: Unequal network configurations, or unequal network returns? *Social Networks*, 70, 138–151.
- Braun, V., & Clarke, V. (2021a). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352.
- Braun, V., & Clarke, V. (2021b). *Thematic analysis: A practical guide*. SAGE.
- Callander, J. K., Johnson, D. E., & Grandis, J. R. (2025). Social capital as a catalyst for gender inequality: A scoping review of networking disparities in academic medicine. *Cancer*, 131(1), e35667.
- Demirci, J. R. (2024). About research: Conducting better qualitative interviews. *Journal of Human Lactation*, 40(1), 21–24.
- Devi, S., Thinakaran, R., Hanefar, S. B. M., & Nadzri, N. R. M. (2024). Tracking academic contributions to women's empowerment in Malaysia: A bibliometric investigation. *Heliyon*, 10(17).
- Eagly, A. H. (2013). *Sex differences in social behavior: A social-role interpretation*. Psychology Press.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598.
- Eaton, A. A., Saunders, J. F., Jacobson, R. K., & West, K. (2020). How gender and race stereotypes impact the advancement of scholars in STEM: Professors' biased evaluations of physics and biology post-doctoral candidates. *Sex Roles*, 82(3), 127–141.
- El-Alayli, A., Hansen-Brown, A. A., & Ceynar, M. (2018). Dancing backwards in high heels: Female professors experience more work demands and special favor requests, particularly from academically entitled students. *Sex Roles*, 79(3–4), 136–150.
- Gaiaschi, C. (2025). Breaking the glass door in academia? Looking at the role of scientific fields and contextual factors in moderating the gender gap in recruitment: Evidence from Italy. *Higher Education*, 1–24.
- Galvin, D. J., Anderson, S. C., Marolf, C. J., Schneider, N. G., & Liebl, A. L. (2024). Comparative analysis of gender disparity in academic positions based on US region and STEM discipline. *PLOS ONE*, 19(3), e0298736.

Guest, G., Namey, E. E., & Mitchell, M. L. (2013). *Collecting qualitative data: A field manual for applied research*. SAGE.

Hamisan@Khair, N. S., Muda, K., & Ismail, F. U. N. (2023). The Perception on Gender Stereotype among Students in Malaysian Public Universities. *International Journal of Academic Research in Business and Social Sciences*, 13(7), 1102 – 1114.

Hands, K., Casey, S., Innes, P., Ashford, T., & Blake, J. (2024). Boundary-less and care-full: Women academics' perception and negotiation of work-life domains during the COVID-19 work from home era at an Australian regional university. *Gender and Education*, 36(3), 248–265.

Heilman, M. E., Caleo, S., & Manzi, F. (2024). Women at work: Pathways from gender stereotypes to gender bias and discrimination. *Annual Review of Organizational Psychology and Organizational Behavior*, 11(1), 165–192.

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523.

Huang, J., Gates, A. J., Sinatra, R., & Barabási, A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117(9), 4609–4616.

Islam, M. A., Hack-Polay, D., Rahman, M., Jantan, A. H., Dal Mas, F., & Kordowicz, M. (2023). Gender and leadership in public higher education in South Asia. *Studies in Higher Education*, 48(8), 1197–1215.

Kariou, A., Koutsimani, P., Montgomery, A., & Lainidi, O. (2021). Emotional labor and burnout among teachers: A systematic review. *International Journal of Environmental Research and Public Health*, 18(23), 12760.

Khokhlova, O., Lamba, N., & Kishore, S. (2023). Evaluating student evaluations: Evidence of gender bias against women in higher education. *Frontiers in Education*, 8, 1158132.

Lean In., & McKinsey & Company. (2022). *Women in the workplace 2022*.

Lim, W. M. (2025). What is qualitative research? *An overview and guidelines*. *Australasian Marketing Journal*, 33(2), 199–229.

Llorens, A., Tzovara, A., Bellier, L., Bhaya-Grossman, I., Bidet-Caulet, A., Chang, W. K., ... & Dronkers, N. F. (2021). Gender bias in academia: A lifetime problem that needs solutions. *Neuron*, 109(13), 2047–2074.

Manzi, F., Caleo, S., & Heilman, M. E. (2024). Unfit or disliked: How descriptive and prescriptive gender stereotypes lead to discrimination against women. *Current Opinion in Psychology*, 60, 101928.

Meza-Mejia, M. D. C., Villarreal-García, M. A., & Ortega-Barba, C. F. (2023). Women and leadership in higher education: A systematic review. *Social Sciences*, 12(10), 555.

Ministry of Higher Education Malaysia (MOHE). (2022). *Malaysian Ministry of Higher Education report: Academic staff demographics in 2021–2022*. MOHE.

Misra, J., Kuvaeva, A., O'Meara, K., Culpepper, D. K., & Jaeger, A. (2021). Gendered and racialized perceptions of faculty workloads. *Gender & Society*, 35(3), 358–394.

- Morgan, H. (2022). Conducting a qualitative document analysis. *The Qualitative Report*, 27(1), 64–77.
- Moorthy, K., Salleh, N. M. Z. N., T'ing, L. C., Ling, L. P., Min Yeng, D., Jia Ning, L., ... & Pui Mun, L. (2022). Gender inequality affecting women's career progression in Malaysia. *Journal of International Women's Studies*, 23(1), 310–332.
- Munusamy, S., Ramasamy, S., & Sukir, N. I. (2024). A systematic review on mental health among educators in Malaysia. *BMC Public Health*, 24, 2634.
- Norander, S., & Zenk, L. (2023). The invisible labor for emerging women leaders. *Advancing Women in Leadership Journal*, 42, 12–22.
- Nurumal, M. S., Zahir, N. M., Ridzwan, M. D. F. M., Hasan, M. K. C., & Sutrisno, R. Y. (2023). Experiences and challenges among female academics during COVID-19. *International Journal of Care Scholars*, 6(1), 92–118.
- Nursyamsi, I., Hidayat, D. R., Melmambessy, D., Ariadi, W., & Wahyudin, M. N. (2024). Trends in research on female lecturers: A bibliometric analysis. *ICE-BEES-24*, 338–357.
- OECD. (2024). *The state of academic careers in OECD countries: An evidence review*. <https://doi.org/10.1787/ea9d3108-en>
- Oliveira, D. F., & Huang, Q. (2025). *Navigating gender disparities in communication research leadership*. Annals of the International Communication Association.
- Othman, N., Rahman, R. S. A. R. A., Abdullah, N., & Ramly, R. (2023). Female leadership in Malaysia: Issues and challenges. *Asian Women*, 39(1), 1–21.
- Peng, H., Teplitskiy, M., Romero, D. M., & Horvát, E. Á. (2025). The gender gap in scholarly self-promotion on social media. *Nature Communications*, 16(1), 5552.
- Publications Office of European Union. (2021). *She figures 2021: Gender in research and innovation*.
- Reid, R. A. (2021). Retaining women faculty: The problem of invisible labor. *PS: Political Science & Politics*, 54(3), 504–506.
- Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F., & Syam, H. (2022). Semi-structured interview: Methodological reflection. *IOSR-JRME*, 12(1), 22–29.
- Sarjidan, M. A. M., & Kasim, A. M. (2023). Trends of academic publications: A case study of Malaysian research universities. *Journal of Research Management and Governance*, 5(1), 18–29.
- Smith, S., & Walker, D. (2024). *The instrumental academic*. *Higher Education Quarterly*, 78(4), e12551.
- Steinþórsdóttir, F. S., Carmichael, F., & Taylor, S. (2021). Gendered workload allocation in universities. *Gender, Work & Organization*, 28(5), 1859–1875.
- Sümer, S., & Eslen-Ziya, H. (2023). Academic women's voices on gendered divisions of work. *European Journal of Women's Studies*, 30(1), 49–65.
- Taparia, M., & Lenka, U. (2022). An integrated conceptual framework of the glass ceiling effect. *Journal of Organizational Effectiveness*, 9(3), 372–400.

Wei, Y., Subramaniam, G., & Wang, X. (2025). The glass ceiling perception and female teacher burnout. *Frontiers in Psychology, 16*, 1551903.

World Bank. (2021). *Malaysia Country Gender Note 2021*.

Appendix A: Semi-Structured Interview Guide

This semi-structured interview guide provided a flexible framework for the in-depth interviews. Interviews were conducted in a conversational manner, allowing the sequence and depth of questions to be adapted to participants' responses. Probing questions were used selectively to elicit clarification or elaboration. Core questions are mapped to the relevant research questions (RQ1 and RQ2). Minor revisions were made following a pilot interview to enhance conceptual clarity and contextual relevance.

A.1 Interview Opening and Informed Consent

Introduction of the researcher and overview of the study purpose;

Explanation of interview procedures, including audio recording, confidentiality, voluntary participation, and the right to withdraw at any time:

Obtaining written informed consent;

Opportunity for participants to raise preliminary questions.

A.2 Professional and Academic Background

Primary Question:

Could you briefly describe your professional background as an academic in a Malaysian public university?

Indicative Probes:

How long have you been working within the Malaysian public university system?

What is your primary academic discipline or area of teaching and research?

What is your current academic position or rank?

Purpose:

To establish contextual background and rapport.

A.3 Experiences of Gender Stereotypes in Academia (Mapped to RQ1)

Primary Exploratory Question:

Based on your experience, how would you describe the gender stereotypes or gender-related expectations that exist within the academic profession?

(Pilot revision: The original wording referred specifically to the “teaching profession.” This was revised to “academic profession” to better reflect the broader scope of university roles, including research, service, and leadership.)

Indicative Probes:

In what ways do these stereotypes manifest in your daily academic work or institutional environment?

How do such stereotypes shape expectations regard academic roles or responsibilities for women and men?

Are certain academic tasks, subjects, teaching styles, or service roles perceived as more suitable for one gender?

Can you share specific examples of gender-based expectations you have encountered, either personally or through colleagues?

A.4 Perceived Impact on Work Performance and Career Development (Mapped to RQ2)

Primary Exploratory Question:

Reflecting on your own career, how do you think gender stereotypes have influenced your professional experiences and career development?

Indicative Probes:

Have gender stereotypes affected your access to opportunities such as leadership roles, research projects, promotion, or professional networks?

How, if at all, have these experiences influenced your job satisfaction, motivation, or sense of professional fulfilment?

Have you experienced challenges in career progression that you believe are related to gendered expectations?

How would you describe the broader gender dynamics in your department or institution, and their influence on professional development?

A.5 Perspectives on Gender Equality and Institutional Change

Primary Question:

From your perspective, how would you describe the current state of gender equality within Malaysian public universities?

Indicative Probes:

Have you observed any progress or initiatives aimed at addressing gender stereotypes or inequality?

What institutional, cultural, or policy changes would you most like to see to improve gender equity?

How could the university system better support gender equality and the professional development of academics?

A.6 Interview Closure

Invitation for participants to share any additional thoughts or experiences not previously discussed;

Expression of appreciation for participation and contributions;

Brief explanation of subsequent steps (e.g., transcription and data handling procedures).

Appendix B: Coding Procedures and Codebook Excerpt

This appendix provides transparency regarding the qualitative data analysis process. It outlines the coding procedures employed in the study and presents an excerpt of the final codebook, illustrating the four core themes reported in the findings.

B.1 Coding Procedures

Interview transcripts were analysed using thematic analysis following Braun and Clarke's (2021) approach. Analysis involved repeated reading of the transcripts to achieve familiarisation, followed by inductive generation of initial codes grounded in participants' accounts. Codes were iteratively refined through constant comparison across interviews, with emerging patterns progressively organised into candidate themes. Analytic memos were maintained throughout the process to document coding decisions, reflexive insights, and theme development. Ongoing discussion among the research team supported clarification of interpretations and consolidation of the analytic framework.

Two coders as main coders of our team were involved in the analysis: the first author and a trained research assistant with prior experience in qualitative coding.

As this study adopted a reflexive thematic analysis approach, formal inter-coder reliability statistics were not calculated. Instead, analytic rigor was supported through investigator documentary contextualisation, independent coding by multiple researchers, reflexive discussion to resolve discrepancies, and the development of a shared codebook.

B2. Codebook Excerpt:

Code Label	Operational Definitions	Inclusion Criteria	Exclusion Criteria	Exemplar Quotes	Participant ID	Analytic Memo
Emotional labour expectation	Expectations that female academics should provide emotional support, counselling, and pastoral care to students	Mentions of being expected to handle student emotions, counselling, or pastoral care without formal recognition	Mentions of formal teaching duties or research responsibilities	"I don't know... it just feels like, somehow, there's this unspoken expectation that... female academics are the ones who should provide emotional support to students, you know?"	P3	This code captured repeated references to uncompensated student care. When grouped with similar codes (e.g., empathy expectations), it pointed to the invisible burden theme.
Overproving Competence	The need to work harder than male colleagues to gain the same level of recognition or credibility	Descriptions of having to constantly prove oneself, overprepare, or exceed expectations to be taken seriously	General mentions of hard work not linked to gender comparison	"Honestly... I've had to keep proving myself again and again... like, working harder than some of my male colleagues just to get the same recognition... it's kind of exhausting."	P1	Participant described having to work harder than male colleagues to gain recognition. This code revealed a higher-order theme of constant need to prove competence as a gendered career barrier.
Authority Questioned	Experiences where professional authority or expertise is doubted or challenged due to gender	Mentions of being interrupted, ignored, or having ideas dismissed in meetings; students questioning authority	General leadership challenges not attributed to gender	"I remember during a departmental meeting... I mean, we had very similar ideas, right? But somehow... his was taken more seriously than mine... I was just sitting there thinking, 'okay... why?'"	P2	This code captured how gender undermines professional credibility. When grouped with "Student Perception Bias," it revealed a broader theme of authority perception disparities, showing that gendered authority gaps operate in both colleague interactions and

						classroom settings.
Promotion Bias	Perceived gender-based barriers to career advancement and leadership positions	Mentions of being overlooked for promotion, leadership roles, or research opportunities due to gender	General career dissatisfaction not linked to gender	"Well... I feel like women are somehow overlooked when it comes to leadership roles... I mean some people just assume we're not as capable as others(male colleagues)..."	P3	This code linked individual experiences of bias to tangible career outcomes. Together with "Authority Questioned" and "Gendered Task Allocation," it contributed to the higher-order theme of constant need to prove competence and career advancement barriers.
Student Perception Bias	Gendered expectations from students regarding teaching style, authority, and approachability	Mentions of students expecting female teachers to be more patient/nurturing and male teachers more assertive/disciplinary	General teaching style preferences not linked to gender	"You know, some students may have these expectations... like, they expect ... it's like we're supposed to be patient all the time. I totally understand them but why always us (female)?"	P3	This code extended the authority theme beyond peer interactions into the classroom. It showed that credibility gaps are not only imposed by colleagues but also reinforced by students' gendered expectations, strengthening the theme of authority perception disparities as a multi-layered phenomenon.
Gendered Task Allocation	Unequal distribution of administrative, teaching, and service tasks based on gender	Mentions of women being assigned more administrative/teaching/service work; men assigned	General workload complaints not linked to gender	"Sometimes I feel like... a lot of the administrative work and committee stuff just automatically comes to us... while my male colleagues can focus more on their research..."	P2	This code showed how gendered expectations translate into concrete workload inequalities, directly feeding into "Promotion Bias" (less time for research) and "Emotional

		more research/leadership tasks		it's a bit unfair, honestly."		Labor" (invisible service work). It contributed to the theme of gendered task distribution and inequalities in workload.

Appendix C: Documentary Sources Used for Analysis

Document title	Source	Year	Analytic role
Malaysian Ministry of Higher Education report: Academic staff demographics in 2021–2022.	The Web of Ministry of Higher Education Malaysia (Open Source)	2022	Provided national-level contextual data on gender distribution across academic ranks and senior positions in Malaysian public universities.
She figures 2021: Gender in research and innovation.	The Web of Publications Office of the European Union (Open Source)	2021	Used to contextualise gender disparities in academic careers and research participation, informing interpretation of participants' accounts of inequality
Women in the workplace 2022.	The Web of Lean In (Open Source)	2022	Provided a comparative organisational context on gendered workload

			expectations, career barriers, and gender inequality trends, serving as a broader reference point for interpreting participants' reported experiences.
Malaysia Country Gender Note 2021.	The Web of World Bank (Open Source)	2021	Used to contextualise broader structural and institutional gender inequalities in Malaysia, providing national-level background relevant to understanding gendered experiences within higher education.

Appendix D: Excerpts of Code-emergence Matrix (From Initial Codes to Core Themes)

Initial Codes	Focused Codes	Core Themes
Emotional support, counselling students, pastoral care, empathy expectations	Emotional labour expectations	Emotional Labor Expectations and the Invisible Burden
Administrative work, committee coordination, event organisation, service roles	Gendered task allocation	Gendered Task Distribution and Inequalities in Workload

Being interrupted, ideas ignored, classroom discipline issues, leadership credibility	Authority perception bias	Authority Perception Disparities and the Credibility Gap
Working harder, higher evaluation standards, delayed promotion, self-doubt	Prove-it-again pressure	Constant Need to Prove Competence and Career Advancement Barriers

Appendix E: Code-Accumulation Matrix.

Interview	New Codes Identified	Cumulative Total
Interview 1 (P1)	12	12
Interview 1 (P2)	4	16
Interview 1 (P3)	1	17
Interview 1 (P4)	0	17

Adlerian-Based Strategies to Support Students' Self-Efficacy and Higher Education Readiness: A Narrative Review

ABSTRACT

Authors: Husna Adilah Mahadi^{1,2}, Noor Syamilah Zakaria^{1*}, Dharatun Nissa Puad Mohd Kari¹

Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

²Faculty of Business, Economics & Social Development, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

*Corresponding author:
syamilah@upm.edu.my

Received: 31/12/2025 | **First Revision:** 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Higher education institutions play a central role in fostering sustainable and inclusive learning environments, yet many students struggle during the transition from school to university due to increased academic expectations, emotional strain, and social adjustment challenges. Self-efficacy has been widely recognized as a key psychological resource that supports students' confidence, motivation, and persistence throughout this transition. However, despite growing interest in counselling interventions that enhance self-efficacy, limited attention has been given to the contribution of Adlerian principles within educational contexts. This gap restricts the integration of theoretically grounded, humanistic approaches in supporting student readiness and long-term academic sustainability. This narrative review therefore aims to synthesize existing literature on Adlerian concepts applied to strengthen students' self-efficacy in educational settings. A total of seven empirical and conceptual studies were analyzed. These studies were published between 2014 and 2025 and were conducted across diverse geographical contexts, including Iran, Indonesia, South Korea, Pakistan, and the United States. Five core themes emerged: (i) encouragement as a fundamental mechanism, (ii) reducing inferiority feelings and fostering courage, (iii) promoting social interest and interpersonal connection, (iv) purposeful behaviour and goal setting, and (v) recognizing personal strengths to build a positive self-concept. Collectively, these themes demonstrate Adlerian counselling's capacity to enhance students' cognitive, emotional, and motivational functioning, all of which underpin academic self-efficacy. The review recommends incorporating Adlerian-informed strategies into educational counselling, transition programs, and student support initiatives to strengthen learners' confidence, resilience, and adaptability which are the key components for sustainable success in higher education.

Keywords: Adlerian counselling, academic self-efficacy, school-to-university transition, encouragement, narrative review

1. Introduction

Higher education institutions are the significant key agents for sustainability development, as this is where the future generation leaders are educated, empowered and being prepared (Michael et al., 2020). In supporting the implementation of Sustainable Development Goals (SDGs), higher education helps in shaping responsible mindset of these future leaders, producing research and actions that support goals like reducing poverty, promoting health, gender equality, consumption responsibility and protecting the environment (Žalėnienė & Pereira, 2021). As a potential driver of sustainability and the SDGs, the earliest factor to consider is how well students adapt to the transition from school-based education into higher educational learning as their ability to benefits from these efforts depends on this adjustment.

Püttmann and Gessner (2023) elaborate students' perspectives about the challenges in their transition to higher education such as emotional barriers, lack of academic preparedness, economic difficulties and family background. The stress, anxiety, fear and confusion are associated with the concerns of academic expectation, hazing, social and environmental adjustments (Dias & Sá, 2014; Zakaria & Warren, 2016). During this shift, students need to prepare and manage their living responsibility, adjusting to more independent learning and social adjustment which are the contributing factors to their mental health (Thompson et al., 2021; Zakaria et al., 2017). The failure of adapting to this shift might heightened risk of mental health difficulties. In relation to this, students' self-belief or confidence in their competence would be a help for a smoother transition in their adaptation period (Püttmann and Gessner, 2023).

The self-belief in this context is closely related to the concept of self-efficacy which was defined by Bandura (1997) as belief in one's capability to execute actions toward desired goals. It is widely acknowledged that self-efficacy is one of the psychological constructs that influences student's learning journeys (Pajares, 1996). Quality Assurance Agency Scotland (2023) outlines four transition stages to higher education going through by students, and it highlights the role of self-efficacy during these stages. Self-efficacy helps students to prepare confidently to enter higher education in pre-transition stage; managing emotional and behavioural response to stress and uncertainty in shock stage; adapting to new academic, social and personal demands in adjustment stage; and potentially supporting long-term development and continued growth during progression stage.

Recent literature examines the role of self-efficacy in education as the potential determinant of students' learning outcomes. For instance, academic self-efficacy has been found to mediate the relationship between study habit and procrastination (Svartdal et al., 2022); predictor to academic achievement (Fakhrou & Habib, 2021); predicts critical thinking (Vachova et al., 2023) and positively correlate with academic competitiveness and cognitive flexibility (Zayed, 2024). Likewise, both academic self-efficacy and general self-efficacy, along with life satisfaction were shown to predict student's perception of a happy school (Döş, 2023). Conversely, low levels of self-efficacy among students have been associated with elevated risk of school dropout (Pedditzi, 2024). Therefore, Meng and Zhang (2023) suggested that higher education institution should have the initiatives in helping the students feel supported, connected and belong to academic community that positively associated with academic self-efficacy and academic outcomes.

Universities should adopt and design innovative, structures, interactive activities that enhance students' responsibility and autonomy in order to help them enhancing their self-efficacy in learning by targeting both psychological and academic outcomes (Yang et al., 2025). In strengthening students' belief in their academic capabilities, guidance and counselling can be the important mechanisms. According to Mathwasa and Sibanda (2020), guidance and counselling enhance students' self-efficacy by fostering self-awareness, emotional stability and confidence, which in turn their decision making, interpersonal skills, motivation, managing academic and personal challenges, and adaptability toward future demands (Carnes-Holt et al., 2016; Subarimaniyam et al., 2020; Warren et al., 2012; Zakaria et al., 2020). Information services, guidance group and counselling group services when infiltrated using varied

methods and tailored to student's characteristics and needs are among the effective counselling approaches for increasing student's academic self-efficacy (Wahid & Neviyarni, 2024; Zakaria et al., 2022).

The effectiveness of counselling approaches in enhancing self-efficacy has been consistently demonstrated across different domain in learning. Lee et al. (2025) reported a significance improvement in academic self-efficacy and learning motivation among students after group counselling intervention, while Rismawan and Gading (2021) found that behavioural group counselling improve career decision making self-efficacy by modifying their dysfunctional thoughts. While counselling in general has been shown to enhance self-efficacy of students, understanding this matter through the lens of relevant counselling theories would be crucial for providing deeper insight into the underlying processes and mechanisms of change.

Sweeney (2009) described self-efficacy as the ability to reason, solve problems and cope effectively with life tasks, reflecting the Adlerian concepts of self-worth, self-evaluation and encouragement through recognition of abilities, value-based decision-making, and meaningful contribution to others. Adlerian theory posits that human beings are inherently social beings who needs others to survive (Gehart, 2016). This social interest reflects not only a sense of belonging among community, but also the willingness to contribute meaningfully to others. However, Adler (1927) proposed that feelings of inferiority can interfere with an individual's striving for growth and the successful fulfilment of life tasks. Thus, Adlerian principles of social interest, encouragement and purposeful living align closely with the development of self-efficacy, as they emphasize positive self-evaluation through connection with others and goal-directed action. To ensure conceptual clarity, several key Adlerian constructs applied in this review are operationally defined as follow:

Encouragement refers to inspiring or helping others toward a conviction that they can work on finding solutions and that they can cope with any predicament (Sweeney, 2009).

Social Interest refers to a person subjectively experiencing a sense that he or she has something in common with other people, is part of community, and benefits from cooperating with others in the community (Gehart, 2016).

Purposeful behaviour refers to teleological and goal-oriented behaviour, where people strive toward meaningful activity, success and achievement (Gehart, 2016).

Inferiority feelings refer to an individual's belief that they are unable to meet standards set by themselves or others, arising from inherent limitations and serving as the basis for striving toward superiority and a sense of completeness (Sipon & Hussin, 2022).

There are numerous studies on Adlerian concepts in educational environment such as (Badejo & Bola, 2020; Bilqis, 2021; Kim & Lee, 2025). These studies discuss the implication of Adlerian principles in enhancing learning process and outcomes among students. Meanwhile, self-efficacy study also has been popular topic on literature since before and recent, especially on its impact on student's academic (Basileo et al., 2024; Dadandi, 2023; Hamann et al., 2020; Manea et al., 2024). In light of this, numerous intervention-based studies on self-efficacy, both in educational settings and other fields, have been grounded in counselling theories. Existing literature demonstrates range of counselling approaches including Solution-Focused Brief Therapy, career counselling (Stipanovic et al., 2017; Türk & Katmer, 2019; Yu et al., 2024) and Cognitive Behavioural Therapy (CBT) (Budiman et al., 2020; Kaplan et al., 2023; Zeidi et al., 2020) are widely applied in improving one's self-efficacy. These approaches primarily strengthen self-efficacy through cognitive structuring, goal clarification, skills development and motivational support.

However, there has been relatively little interest to apply Adlerian theory as the theoretical framework for enhancing self-efficacy, particularly within educational context. This imbalance in literature highlights a need for greater attention to Adlerian principles, particularly given its emphasis on social

interest, encouragement and purposeful living in enhancing students believe in their capabilities. Given the potential of Adlerian concepts in enhancing self-efficacy, and the limited empirical evidence focusing on this matter, it becomes necessary to synthesize previous literature that have explored Adlerian-based intervention or principles on self-efficacy.

Therefore, this review adopts a theory-building narrative synthesis approach based on small and heterogeneous body of literature ($n=7$) aimed to explore and discuss the available studies on how Adlerian principles could address students' self-efficacy in educational contexts. This paper focused on specific interventions or techniques grounded in Adlerian principles that are designed to enhance student's self-efficacy. Thus, it provides synthesis of relevant concepts and strategies, and their potential impact on students' belief on their capabilities, particularly supporting their transition to higher education. Accordingly, this review potentially offers useful insights for counsellors in designing and facilitating theory-driven interventions that nurture student's self-efficacy, which lead to their adjustment and foster their academic outcomes. Therefore, the aim is not to make general conclusions, but to provide a conceptual understanding of how Adlerian principles may support students' self-efficacy and to suggest practical strategies for educational support.

Research objectives to guide this review are as follow:

1. RO1: To summarize empirical and conceptual studies that apply Adlerian principles to self-efficacy in educational settings.
2. RO2: To identify key Adlerian intervention features that support students during educational transitions.

Based on these objectives, the following research question is derived:

1. RQ1: What Adlerian concepts, principles and techniques have been identified in the literature as contributing to the enhancement of self-efficacy in educational contexts?

2. Methodology

In addressing the research question, this study employed a narrative review approach to synthesize and discuss the related existing evidence. Sukhera (2022) defines narrative review as a form of knowledge synthesis that provide summary, interpretation and critique of findings from broad range of studies in understanding a topic. Unlike systematic review that follow rigid protocol and focus on narrow research questions, narrative review allows critical interpretation and inclusion of diverse study type with more flexible, non-systematic approach.

To justify the selection of narrative review methods for this study, explanation on the strengths and weaknesses was reasonable, as described by Bourhis (2018). Bourhis explains that narrative review's strength lies on its flexibility enabling the researcher to combine different types of research, synthesizing findings in-depth, and placed the studies in wider historical and theoretical context. This approach is suitable for exploration of idea development and correlating various perspectives on the topic. However, narrative review has its own limitation, including vulnerability toward type I and type II in individual study, the absence of method to reduce error, combined with heavy reliance on researcher's interpretation which may increase subjectivity and reduce reproducibility. Therefore, even though narrative review gives more in-depth understanding on theory, the conclusion should be interpreted cautiously due to the potential inconsistencies within the underlying empirical evidence.

2.1. Inclusion and exclusion criteria

Guided by Adlerian theory in counselling, the first step in selecting the relevant studies involved defining the theoretical foundation and specifying the research question, which focused on identifying the Adlerian concepts and techniques applied to enhance self-efficacy. In determining the required

characteristics of primary studies, inclusion and exclusion criteria was decided to ensure only relevant studies included and answer the research question. Table 1 outlines the inclusion and exclusion criteria established for selecting studies. These criteria help in guiding the review process in identifying the relevant literature that applied Adlerian principles within its intervention or discussion related to self-efficacy in educational context.

Table 1. Inclusion and Exclusion Criteria

Criteria Type	Inclusion Criteria	Exclusion Criteria
Type of study	Peer-reviewed journal articles include empirical or theoretical.	Non-scholarly sources or non-peer-reviewed materials and grey literature (eg., theses, reports, conference paper).
Focus area	Studies that discuss or examined self-efficacy, academic self-efficacy, learning related self-efficacy or concepts theoretically linked to self-efficacy within Adlerian framework.	Articles unrelated to self-efficacy or not conceptually connected to it.
Educational context	Studies conducted in educational, school, university, training context, or those involving students.	Studies not involving students' population or non-educational settings.
Theoretical relevance	Studies that explicitly or implicitly applied Adlerian theory, Adlerian counselling principles or concepts (including theoretical, conceptual, or intervention-based work).	Studies grounded in other counselling or psychological frameworks, without reference to Adlerian concepts.
Language	Written and published in English.	Written and published in other than English languages.

To provide a comprehensive understanding of how Adlerian principles contribute to the development of self-efficacy in educational settings, this review included both empirical studies and conceptual papers. The limited number of empirical investigations applying Adlerian approaches directly to self-efficacy necessitated the inclusion of theoretical literature to strengthen the review's foundational perspective. Conceptual papers offered essential insights into core Adlerian principles such as encouragement, social interest, and purposeful behaviour, that help explain the mechanisms through which self-efficacy may be enhanced, leading to better transition to higher educational level. Including both types of literature allowed for a richer and more integrated synthesis of Adlerian concepts relevant to education.

2.2. Search strategies and selection

Retrieving a sample of potentially relevant literature involves choosing the right keywords and effective search strings that could answer the research questions and match the research purpose (Tranfield et al., 2003). The literature search was conducted in two stages. An initial search was carried out between August and September 2025, followed by an updated search conducted up to 20 January 2026 to capture potential newly published and previously missed studies. The literature search used a combination of keywords representing the main constructs for this review, including “Adler”, “Adlerian Counselling”, “Self-efficacy”, “Academic self-efficacy”, and “education”. Boolean operators (AND, OR) were applied to combine selected keywords across structured online database using Scopus, SAGE and ERIC Journal. Additional databases were initially explored; however, no further studies meeting the specific focus on Adlerian-based strategies and self-efficacy outcomes in educational context were identified. In addition, a manual search was also conducted using Google Scholar to identify studies that might not be indexed in major databases, and to ensure the comprehensive coverage of potential studies. Database-specific search strings were adapted according to each platform's indexing system and are presented in Appendix A.

The search was limited to studies published between 2010 and 2026. This timeframe was selected to capture contemporary applications of Adlerian’s counselling interventions and self-efficacy within educational contexts, while avoiding very early studies that may not reflect current theoretical developments and educational practices. Only peer-reviewed articles, written and published in English were included for this review. Following the database and manual searches, records were screened at the title and abstract level to evaluate its relevancy for this review. Then only the full-text screening was conducted to select the potential study based on pre-defined inclusion and exclusion criteria. These processes lead to final seven studies that identified as relevant studies to answer the research question. Recent study (2025) included in this review were published in peer-reviewed journals. Therefore, its findings are considered credible, although they are interpreted with appropriate caution due to their regency. The study selection process is illustrated in Figure 1, a PRISMA-style flow diagram, detailing the identification, screening, eligibility and inclusion stages.

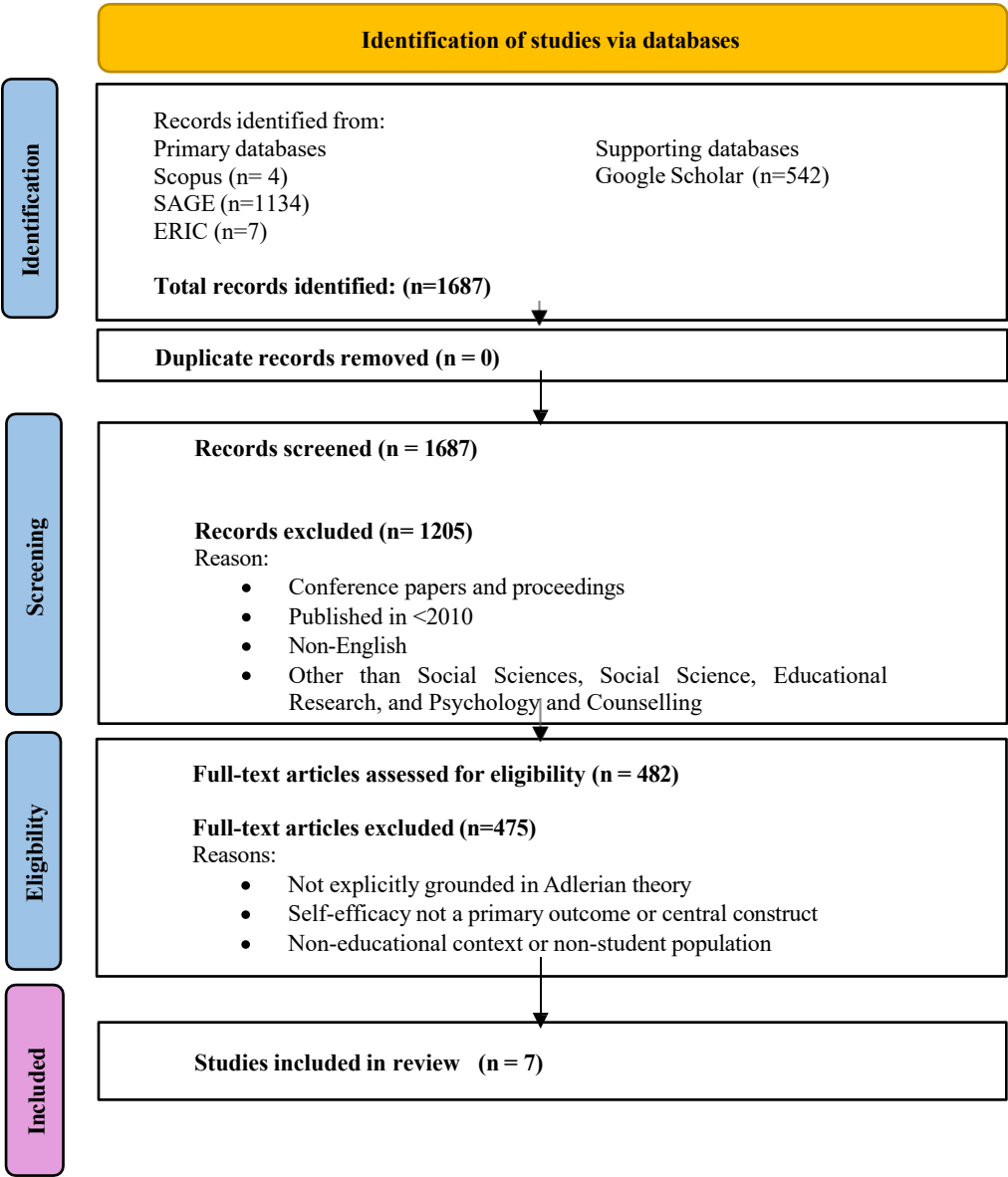


Figure 1. PRISMA-Style Flow Diagram of Study Selection

2.3. Quality appraisal of empirical studies

A brief methodological appraisal was conducted for the empirical studies included in this review. It is to describe the design characteristics and to assess overall study quality. The appraisal focused on the study design, sample size, intervention type and outcome measures that related to self-efficacy. As this study is a narrative review, no formal risk-of-bias scoring is applied. Table 2 outlines the methodological characteristics and descriptive quality appraisal of included empirical studies with the quality rating for each study.

Table 2. Methodological Characteristics and Descriptive Quality Appraisal of Included Empirical Studies

Author, Year	Study Design	Sample Size	Methodological Notes	Quality Rating
Farahbakhsh et al. (2014)	Quasi-experimental, pre-test-post-test with control group (waiting list).	n = 28 teachers.	Convenience sampling; no randomisation at school level; small sample size; outcomes measured indirectly via teacher intervention.	★★
Lee et al. (2025)	Randomized controlled trial with pre-test, post-test, and follow-up (repeated-measures design).	n = 40 university students.	Conducted at a single university; participants were students receiving academic warnings; relatively small sample size; follow-up period limited to one post-intervention assessment.	★★★
Fatima & Ashraf (2018)	Correlational research design with mediation analysis.	n = 440 undergraduate and graduate students.	Cross-sectional design; purposive sampling; reliance on self-report questionnaires; mediation tested using PROCESS macro.	★★
Dortaj & Sohrabi (2020)	Semi-experimental (quasi-experimental) pre-test-post-test design with control group.	n = 60 female high school students.	Multi-stage cluster sampling; participants drawn from a single public high school in Tehran and from female group; outcomes measured at pre-test and post-test only; no follow-up assessment conducted.	★★
Sundah (2019)	Counselling action research (classroom) conducted in iterative stages (planning, action, observation, reflection).	n = 8 middle school students.	Very small sample size; single-school setting; absence of control or comparison group; qualitative descriptive analysis used; outcomes assessed within the action research framework.	★
Safara et. al., (2020)	Quasi-experimental two-group pre-test-post-test design (experimental and control groups).	n = 30 male high school students.	Cluster random sampling from a single high school; participants limited to male students; outcomes assessed at pre-test and post-test only; no follow-up assessment reported.	★★

Note: ★★★=strong (e.g., randomized or well-controlled design); ★★=moderate (e.g., quasi-experimental); ★= limited (e.g., small-scale study)

The methodological appraisal indicates considerable variation in study designs, sample sizes and outcome measures across the empirical studies included in this review. While one randomized controlled trial demonstrated stronger methodological rigor, most studies employed quasi-experimental or correlational designs with relatively small or context-specific samples and limited follow-up assessments. A simple quality appraisal rating was conducted by categorising the included studies as

strong (★★★), moderate (★★) or limited (★) based on the research design, sampling size, and methodological rigor. This indicator informed the interpretation of the findings, with greater emphasis placed on studies with stronger methodological design. Despite its constraints, all empirical studies consistently reported positive associations between Adlerian-based strategies and students' self-efficacy outcomes. In addition to the empirical evidence, one conceptual paper was included to elucidate the theoretical relevance of Adlerian principles such as encouragement, social interest, and purposeful behaviour in supporting the development of self-efficacy among students.

2.4. Data synthesis

To answer the research question related to the identification of counselling strategies based on Adlerian theory in increasing students' self-efficacy, this review uses a combination of deductive and inductive thematic analysis. Deductive analysis is carried out by developing initial categories based on key concepts in Adlerian Therapy such as encouragement, purposeful behaviour and striving for superiority, in line with the deductive analysis defined by Merriam and Tisdell (2016). Simultaneously, inductive thematic analysis allows data-driven codes and themes to emerge within the categories, identifying specific counselling strategies or forms of intervention utilized in selected studies. The initial coding was performed by the main author and iteratively refined through two rounds of review. Codes were then grouped into broader categories and discussed with co-authors in three meetings to reach agreement. Through this iterative process, the final themes were developed and agreed upon the research team in order to enhance clarity and theoretical coherent. The identified themes are then synthesized narratively to explain the role of Adlerian approaches in supporting the increase in students' self-efficacy, specifically in helping them adapt and make successful transitions to higher education.

In brief, the screening and extracting data processes was mainly done by the main author. The title and abstract primarily screened based on predefined inclusion and exclusion criteria, followed by full-text evaluation of the selected studies. To improve the reliability and minimize potential selection-biased, screening decisions and extracted data were reviewed and discussed further with the co-authors. Any discrepancies and uncertainties were dissolved through discussion until a consensus was reached. Besides that, the coding and thematic synthesis were refined iteratively through collaborative review from research team to ensure consistency and theory alignment.

Artificial Intelligence-assisted tools were used for this review to support language editing and clarity during manuscript preparation. It did not contribute to the research design, data selection, or analysis. All conceptual development, inclusion and exclusion decision, data interpretation, synthesis of findings and final content decisions were made by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

3. Findings

3.1. Background of the Studies

Seven selected studies were published between 2014 and 2025, representing a decade of research integrating Adlerian principles in educational settings. These studies were conducted across diverse geographical contexts, including Indonesia, Iran, South Korea, Turkey, and Western regions. The samples primarily comprised middle school, high school, and university students, with study designs ranging from empirical intervention programmes to correlational research and conceptual analyses. This variation in publication year, country, and population provides diverse evidence base for synthesising Adlerian concepts relevant to the enhancement of self-efficacy in education. The summary of synthesis for this review is presented in Table 3.

3.2 Themes Derived from The Reviewed Literature

The Adlerian’s principles, concepts and techniques that were identified from the selected studies were concluded in Table 3. The table comprises the Adlerian’s concepts that were discussed and applied in the studies, that were potentially contribute to self-efficacy enhancement of students particularly in academic context. The findings of the literature synthesis outlined five main themes to answer the research questions, namely: (i) encouragement as the core mechanism for strengthening self-efficacy, (ii) overcoming inferiority and developing courage, (iii) social interest and interpersonal connection in building self-efficacy, (iv) purposeful behaviour and goal setting as drivers of self-efficacy development, (v) recognising strengths and positive self-concept development. These five themes were concluded based on the Adlerian elements that was implemented or discussed in their studies. Figure 2 illustrates the themes of Adlerian concepts related to self-efficacy enhancement emerged from the selected studies.

Figure 2. Thematic Synthesis of Adlerian Concepts Related to Self-Efficacy Enhancement

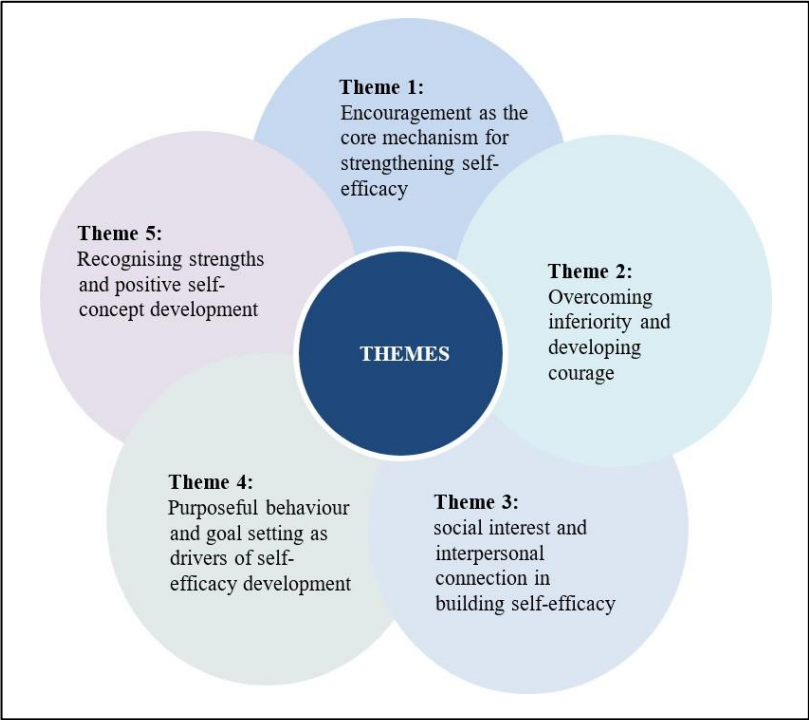


Table 3. Summary of Study Background and Adlerian Concepts and Techniques Used

Author, year, country	Study Design	Sample Size	Population	Type of Intervention	Adlerian Principles, Techniques, or Concepts Identified / Findings
Farahbakhsh et al. (2014) Iran	Quasi-experimental, pre-test-post-test with control group (waiting list).	n = 28 Teachers (14 experimental, 14 control); outcomes measured on students.	Middle school teachers (intervention) and students experiencing academic failure.	Teacher's instruction – Teacher attending several training sessions, learned and practiced Adlerian-based encouragement methods for interacting students with academic failure.	– Train teachers on Adlerian encouragement techniques to replace criticism with positive reinforcement, strengthen social interest, and build supportive teacher-student relationships. – Improvements in students' academic self-efficacy, and other construct such as motivation, emotional regulation and planning.
Lee et al. (2025) Korea	Randomized controlled trial with pre-test, post-test, and follow-up (repeated-measures design).	n = 40 university students (Experimental = 20; Control = 20).	University students under academic warning (South Korea).	Group counselling program – 8-week (120 minutes per session) group counselling program grounded in Adlerian Individual Psychology and Positive Psychology principles.	– The program builds on Adlerian's principle, emphasizing social interest, encouragement, and purposeful living . – Integrated Adlerian concepts and techniques: • Encouragement and social interest (foster belonging and motivation) • Exploration of inferiority and superiority complexes (to help students reframe feelings of failure) • Lifestyle assessment and early recollections (to increase self-awareness and insight into private logic) • Goal setting, "as if" techniques, and strength-based activities (to enhance purpose and responsibility) • Mutual support and community building (contribution to others) – Improved student's academic self-efficacy, and other construct (self-determined learning motivation, GPA).
Safara et al. (2020) Iran	Quasi-experimental two-group pre-test-post-test design (experimental	n = 30 high school students; Experimental = 15, Control = 15.	Male secondary school students (Grade 9) in Iran.	Group guidance session – Self-Encouragement Training Program for 10 sessions (60 minutes each, twice per week for five weeks)	– The intervention employed Adlerian self-encouragement training based on Schwanker's model , incorporating principles such as encouragement, courage to be imperfect, purposeful behavior, positive interpersonal communication, and a sense of belonging. – This helps in:

	and control groups).			– Intervention grounded in Adlerian Individual Psychology following Schwanker’s self-encouragement training model (2009).	• Encouragement leads to greater feelings of competence and capability. • Positive interpersonal interactions improve confidence. • Accepting imperfections reduces fear of failure and supports trying new tasks. • Feeling a sense of belonging improves motivation and belief in one’s abilities.
Fatima & Ashraf (2018) Pakistan	Correlational research design with mediation analysis.	n = 440 students (220 males, 220 females).	Undergraduate and graduate students, aged 17–25, from colleges and universities in Pakistan.	Correlational research study – Used three instruments: the Psychological Birth Order Inventory to assess perceived family position, the Self-Efficacy Scale (Schwarzer & Jerusalem, 1995) to measure beliefs in one’s capabilities, and the Achievement Motivation Scale to evaluate students’ drive to achieve goals.	– Applied the Adlerian concept of psychological birth order to examine its relationship with self-efficacy and achievement motivation among university students. – It aims to explain how family dynamics and perceived roles influence motivational and efficacy beliefs among university students. – It is found that: • Psychological birth order was significantly associated with both variables: first-born students reported higher self-efficacy, while middle-born and only born students showed lower levels, and youngest-born students showed no significant association. • Birth order was also linked to achievement motivation, with first-born and youngest students displaying lower motivation. • Self-efficacy was found to mediate the relationship between certain birth-order positions (middle and only born) and achievement motivation.
Dortaj & Sohrabi (2020) Iran	Semi-experimental (quasi-experimental) pre-test-post-test design with control group.	n = 60 female high school students (Experimental = 30; Control = 30).	Female high school students in public high school in Tehran, Iran.	Training program – Self-encouragement training based on Adler’s Individual Psychology, delivered as an 8-session training program (once a week, 90 minutes per session).	– The intervention applied Adlerian self-encouragement training, which focused on helping students recognise their personal strengths, increase their self-esteem and confidence, build optimism and hope, develop social interest, feel valued and capable, and reduce discouragement – The Adlerian concepts of encouragement: • Encouragement strengthens the internal belief of “I can.”

					• Recognizing strengths increases students' sense of capability. • Developing optimism and hope enhances psychological capital, which includes self-efficacy. • Feeling valued and connected improves motivation and confidence. – Increased the academic engagement and psychological capital of female students include self-efficacy (by helping them recognize their strengths, increase confidence, feel more hopeful and optimistic, and foster social interest).
Sundah (2019) Indonesia	Counselling action research (classroom) conducted in iterative stages (planning, action, observation, reflection).	n = 8 middle school.	Students with low self-efficacy in a middle school in Indonesia.	Group counselling session – Six counselling sessions grounded in Adler's Individual Psychology, which each session targeted different Adlerian theme.	– Intervention used Adlerian Group Counselling, built on the core concepts of Adler's Individual Psychology as session components: • Developing courage (to build students' courage to face academic tasks and challenges) • Encouragement (a central technique used to support students in building confidence and self-belief) • Social interest (students were guided to develop cooperation, belonging, and positive interaction with peers) • Goal setting or purposeful behaviour (help students identify goals and understand purposeful direction in academic life) • Understanding strengths and abilities (students learned to recognise their own potentials and personal abilities).
Wong (2015) USA	Conceptual/theoretical review.	Not applicable.	Not applicable (conceptual paper).	Not applicable Theoretical analysis – presents a comprehensive theoretical analysis of the concept of encouragement, highlighting its Adlerian origins and its relevance across contemporary counselling and psychological practice.	– Encouragement strengthens an individual's belief in their capability, promotes resilience, and reduces feelings of inferiority, aligning closely with Adler's goals of fostering courage and social interest. – Proposed The Tripartite Encouragement Model (TEM) conceptualises encouragement across three domains: foci, features, and levels. The foci distinguish between challenge-focused and potential-focused encouragement. The features describe how encouragement is communicated, emphasising framing, credibility, and reinforcing the individual's

					ability and effort. The levels illustrate that encouragement can occur interpersonally, intrapersonal, or as a group norm. Together, these components explain how encouragement strengthens confidence, capability beliefs, and resilience--core elements relevant to self-efficacy.

4. Discussion

The synthesis of the selected studies highlights multiple Adlerian elements that contribute to development or enhancement of one's self-belief on their capabilities to complete tasks in educational context. These themes demonstrate that Adlerian counselling offers both cognitive and emotional mechanism in helping students to improve their self-efficacy. This were done through supporting one's emotion through encouragement and positive relationship, and changing their though pattern through purposive goal, courage and awareness on self-strength. The discussion below interprets the five emerged themes in relation to Adlerian theory and broader self-efficacy framework.

4.1. Theme 1: Encouragement as the Core Mechanism for Strengthening Self-Efficacy

The prominence of encouragement across all seven reviewed studies reinforces Adler's view that encouragement is a primary mechanism for fostering psychological growth and strengthening capability beliefs. Empirically, encouragement was implemented through positive teacher reinforcement (Farahbakhsh et al., 2014), peer-supported group counselling (Lee et al., 2025; Sundah, 2019), and structured self-encouragement training designed to build courage and self-acceptance (Safara et al., 2020; Dortaj & Sohrabi, 2020). From these studies, two quasi-experimental studies involving secondary school students ($n = 30$ and $n = 60$, respectively) reported significant pre-post improvements in academic self-efficacy following Adlerian-based self-encouragement training programmes. These interventions typically emphasised positive reinforcement, reframing of perceived failure and strengthening students' beliefs in their academic capabilities. In addition, a randomised controlled trial conducted among university students under academic warning demonstrated that group counselling incorporating encouragement-based Adlerian principles led to significant gains in academic self-efficacy compared with a control group. However, across the reviewed literature, empirical evidence remains limited by relatively small sample sizes and short-term outcome assessments, and no large-scale or longitudinal trials were identified. Another interesting finding from Fatima and Ashraf (2018) who discuss this pattern by demonstrating that differing levels of encouragement within family birth-order roles shape students perceived competence, suggesting that encouragement influences self-efficacy even outside formal intervention settings.

These findings are in line with other study, where Keller and Szakál (2021) in their experimental study demonstrated that self-efficacy increases in response to encouraging messages that affirm past success and convey confidence in future performance of students. Its highlight encouragement's impact on cognitive and affective. Taken together, these external and reviewed studies suggest that encouragement plays a dual and important role as it reduces discouragement and simultaneously support capability beliefs, making it a central driver of self-efficacy development in educational contexts.

4.2. Theme 2: Overcoming Inferiority and Developing Courage

The reviewed studies consistently demonstrate that Adlerian interventions help students reduce feelings of inferiority and build the courage needed to engage with academic tasks. Adler conceptualized inferiority as the feeling of incompleteness, helplessness and dependency shaped from childhood that urge the individual to compensate this through guiding fantasies of strength (Adler, 1927). When these feelings become stronger and prolong, it developed to inferiority complex, where individual felt not adequate, perceives others as more capable and feels unable to face challenges. In contrast, superiority complex refers to overly high opinion on oneself. From the review, developing courage is one way to compensate this. Farahbakhsh et al. (2014) showed that shifting from criticism to supportive communication reduces inadequacy and strengthens students' confidence. Similarly, Lee et al. (2025) and Sundah (2019) used group counselling processes such as peer affirmation, shared problem-solving, and reflective dialogue to help students confront discouraging self-beliefs and develop academic courage. Self-encouragement programmes implemented by Safara et al. (2020) and Dortaj and Sohrabi (2020) further reinforced the courage to be imperfect, encouraging students to accept personal limitations while reframing perceived weaknesses more constructively.

Overall, empirical support for this theme is drawn from a small number of intervention studies employing quasi-experimental, action research and randomised controlled designs, with sample sizes ranging from small group interventions to moderate-sized school and university samples. Across these studies, interventions targeting inferiority-related beliefs through encouragement and supportive group processes consistently reported improvements in students' academic self-efficacy and courage to engage with academic tasks, although the evidence remains limited by short-term assessments and modest sample sizes.

These findings align with broader evidence showing that addressing inferiority and promoting courage directly enhances self-efficacy. For example, Berezyuk et al. (2023) found that individuals who acknowledged and worked through inferiority tendencies showed higher self-efficacy than those with superiority complexes, suggesting that confronting inferiority can be an adaptive pathway toward self-belief. Likewise, Habsy et al., (2025) demonstrated that Adlerian techniques in counselling effectively increase student self-esteem by changing student's self-perception which consequently enhance their confidence and courage, while improving their social engagement. Addressing inferiority feelings and cultivating courage, which are one of the potential pathways to strengthening students' academic self-efficacy are highly relevant to students who are entering higher education, where feelings of inadequacy and fear of failures during this phase can hinder their academic and social adjustment (Thompson et al., 2021).

4.3 Theme 3: Social Interest and Interpersonal Connection in Building Self-Efficacy

A few studies explicitly applied Adlerian principles of social interest and interpersonal connection to strengthen students' self-efficacy. Social interest refers to an individual's sense of connectedness with others and genuine concern for the well-being of the broader community and expressed through cooperative behaviour aiming at contributing to collective progress rather than pursuing personal advantage (Gehart, 2016). These concepts were found implemented in the selected empirical studies such as, Farahbakhsh et al. (2014) promoted social interest by fostering supportive teacher–student relationships that enhanced students' sense of belonging. Lee et al. (2025) and Sundah (2019) incorporated peer cooperation and mutual encouragement in group counselling, enabling students to experience acceptance, connection, and shared competence. Wong (2015) also highlighted social interest as a core Adlerian principle, emphasising that feeling connected and contributing to others is foundational to developing confidence and self-efficacy.

The findings of this review align with broader evidence showing that social connectedness and belonging, a central element of Adlerian social interest play a significant role in strengthening self-efficacy. Datu and Yuen (2020) in their correlational study involving high school students in Hong Kong found that strong connections with parents and peers significantly predicted students' academic and career self-efficacy, with gratitude mediating this relationship. Although not framed within Adlerian theory, this reinforces Adler's assumption that a sense of relational embeddedness nurtures psychological resources that support capability beliefs. Similarly, another correlational study with 245 preservice teachers in three university teacher education programmes by Bjorklund et al. (2020) demonstrated that preservice teachers with a stronger sense of belonging and higher network centrality reported higher self-efficacy, supporting the idea that interpersonal connection enhances one's confidence to handle academic or professional demands. This convergence of Adlerian theory and contemporary empirical evidence indicates that fostering interpersonal connection could be meaningful pathway for enhancing self-efficacy for the students, especially during transition period as Püttmann and Gessner (2023) concludes family support and peer relationship were among the most significant facilitators throughout the transition journey. The feeling of belonging and togetherness would help them lessen their burden in this matter.

4.4. Theme 4: Purposeful Behaviour and Goal Setting as Drivers of Self-Efficacy Development

Several studies in this review applied Adlerian ideas of purposeful behaviour and goal-directed action as mechanisms for strengthening self-efficacy primarily employing quasi-experimental and action research designs with modest sample sizes. Safara et al. (2020) and Dortaj and Sohrabi (2020) used self-encouragement training that guided students to set meaningful goals, act intentionally, and monitor their progress, a process that helped them develop a clearer sense of competence. Likewise, Sundah (2019) showed that engaging students in purposeful actions during group counselling reinforced their academic confidence through deliberate effort and persistence. Across these studies, purposeful behaviour and goal-directed activities were embedded within broader Adlerian self-encouragement and group counselling interventions rather than examined as isolated mechanisms. Despite this limitation, consistent pre-post improvements in academic self-efficacy were reported, suggesting that purposeful action may play a meaningful contributory role in strengthening students' self-beliefs. Wong (2015) in his conceptual paper further emphasized that purposeful striving is central in Adlerian counselling, as it directs individuals toward constructive action that nurtures belief in their abilities.

These findings are supported by broader literature linking goal-directed behaviour to self-efficacy development. Sides and Cuevas (2020) demonstrated that receiving feedback on progress toward personal goals enhances students' self-efficacy by providing concrete evidence of capability and improvement. Similarly, Jun (2021) found that students with higher self-efficacy tend to adopt stronger goal orientations and aspire toward more challenging academic and career pathways, underscoring the reciprocal relationship between purposeful striving and belief in one's competence. These studies affirm that goal setting, purposeful action, and progress monitoring, core Adlerian ideas could serve as meaningful pathways for enhancing students' academic self-efficacy. A study in Malaysia reported that self-efficacy concurrently has strong influence on the way university students set their goals and persistence, exhibit healthy self-development growth, and supports important soft skills especially needed for employability (Abdullah et al., 2014). This is among the ground focus that need to be paid attention to when students are preparing another phase of learning in their life.

4.5. Theme 5: Recognizing Strengths and Positive Self-Concept Development

The reviewed studies highlight that recognising personal strengths and cultivating a positive self-concept are central Adlerian strategies for enhancing self-efficacy. Safara et al. (2020) and Dortaj and Sohrabi (2020) demonstrated that self-encouragement training enables students to identify their abilities and replace discouraging self-evaluations with more constructive beliefs. Similarly, Sundah (2019) showed that supportive feedback within Adlerian group counselling helped students acknowledge their competencies, while Wong (2015) emphasised theoretically that affirming strengths is foundational to the development of confidence. The evidence from this review demonstrates that helping students recognise their strengths is essential in building the confidence needed to enhance their self-efficacy.

These past literatures have widely discussed the psychological and academic value of recognising one's strengths. García-Álvarez et al. (2024) found that acknowledging character strengths such as leadership, hope, and persistence significantly predict both general and academic self-efficacy among university students. Recognizing personal strengths is strongly associated with academic achievement, suggesting that students who perceive themselves as capable are more likely to perform well academically (Alipour, 2024). Additionally, Firmansyah (2025) noted that strong self-recognition in self-efficacy encourages the use of metacognitive strategies, which enhances academic performance while fostering resilience and adaptability. Collectively, these studies support the Adlerian view that affirming strengths and developing a positive self-concept provide a psychological foundation upon which stronger academic self-efficacy can be built. This is also a crucial part as recognition in one's strength helps students to feel more confident and quickly adapt to academic demands and perseverance in handling the shift into higher education that are more challenging (Lin et al., 2021).

Taken together, while the findings across the reviewed studies consistently suggested positive relationship between Adlerian-based strategies and students' self-efficacy, a few limitations among this

empirical evidence should be acknowledged. Most of the empirical research employed either quasi-experimental studies (Dortaj & Sohrabi, 2020; Farahbakhsh et al., 2014; Safara et al., 2020) or action research with small sample size (Sundah, 2019), which might limit the generalizability of the findings. Besides, lack of long-term follow-up assessments restricts the ability to determine sustainability of the outcomes that were reported from time to time. Some of the research also put the focus on specific populations such as single-gender samples or students from single institutions (Dortaj & Sohrabi, 2020; Safara et al., 2020), which might limit applicability of the findings to a broader educational contexts. These methodological variations and constraints highlight the need to interpret the findings cautiously and support the positioning of this review as exploratory rather than conclusive.

4.6. Contextual and Cultural Moderators

Another consideration in interpretation of the findings of this review is on sociocultural orientation of the population involved. This study primarily involves countries such as Iran, South Korean, Indonesia and Pakistan which often characterised by more collectivist cultural orientation rather than individualist. As refers to Hofstede (1980), collectivist culture emphasis connectedness, similarities and harmony within their group, whereas individualist want to be autonomous, distinct and independent from others. So, in such context, Adlerian principles such as social interest, encouragement and belongingness may align closely to population from collectivist cultural values.

The cultural factor such as family, teacher and peers' encouragement might influence more to this culture and impact on how they respond to counselling interventions. Therefore, while suggesting that Adlerian-based strategies may support student's self-efficacy, their effectiveness may also vary across cultural and educational context. This highlights the importance of adapting Adlerian-based approach to suit sociocultural background of the student when applied in diverse setting.

4.7. Implications for Higher Education Transition Support Practices

Based on the findings and discussion of this review, the authors propose a brief Adlerian-based strategies according to their level of implementation feasibility and resource requirements: immediate (low-cost), medium-term (moderate cost), and long-term (high cost). Immediate strategies, such as encouragement and peer support, can be applied directly in classroom settings. Medium-term strategies, including small-group counselling and structured programmes, require planning and facilitator training. Long-term strategies, such as curriculum integration and institutional interventions, involve broader systemic changes. This classification provides practical guidance for prioritising implementation based on available resources. The proposed practical implementation is outlined in Table 4.

Table 4. Practical Implementation of Adlerian-Based Strategies

Activity	Feasibility/ Cost
Encouragement practices, positive feedback, peer support activities	Immediate (Low-cost) – Can be implemented directly by lecturers or counsellors in classroom settings to foster students' confidence and sense of belonging
Small-group counselling, structured psychoeducational sessions, mentorship programmes	Medium-term (Moderate cost) – Require some planning and training to support targeted development of self-efficacy through guided activities
Curriculum integration, institutional counselling programmes, policy-level interventions	Long-term (High cost) – Involve systemic changes to embed Adlerian principles into educational practices and student support systems

5. Conclusion

This narrative review synthesised seven studies that applied Adlerian concepts within educational contexts to strengthen students' self-efficacy. The findings demonstrate that Adlerian principles particularly encouragement, overcoming inferiority, social interest, purposeful behaviour, and recognition of strengths play meaningful roles in shaping students' beliefs in their academic capabilities. Across diverse countries, populations, and intervention formats, these principles consistently supported positive cognitive, emotional, and motivational outcomes, including enhanced academic self-efficacy, improved motivation, greater psychological capital, and stronger engagement. Collectively, the review highlights that Adlerian counselling offers a holistic, theory-driven framework capable of addressing students' self-belief through both interpersonal connection and intrapersonal change. This synthesis therefore underscores the potential value of integrating Adlerian-informed strategies into educational counselling practices to support students' learning and development.

Importantly, the findings from this review are closely related to the psychological demands that need to be faced by students during transitional phase from school-based learning to higher education. This phase is typically marked by emotional pressure, new academic expectation, and a more complex adaptation challenges (Dias & Sá, 2014; Thompson et al., 2021). Since self-efficacy plays a significant role in determining how far the students could navigate these changes comprising confidence, persistence and their preparedness in fulfilling universities' demands (Lin et al., 2021; Quality Assurance Agency Scotland (2023), Adlerian mechanisms were identified through this review offers potential pathway in supporting the transition. The elements of encouragement, overcoming inferiority and developing courage, social interest and interpersonal connection, purposeful behaviour and goal setting, and strengths acknowledgement sustain motivation throughout this phase. By strengthening students' sense of abilities and belonging (Abdullah et al., 2014; Meng & Zhang, 2023), Adlerian based strategies potentially playing pivotal role in helping the students in shifting to higher education smoothly, with resilience and confident.

Integrating these psychological principles into higher educational practice are significant in ensuring the potential long-term benefits in educational system. However, this remains to be confirmed through longitudinal studies. As the institutions strive to enhance student's well-being and cultivating resilience learners, strengthening self-efficacy could be the grounding strategies in supporting these goals. Adlerian-based approaches offer a humanistic and theory-driven pathway in helping higher educational institution to foster students who are not only capable in academic, but also socially connected, emotionally prepared and motivated. Therefore, supporting self-efficacy development efforts could potentially contribute significantly to development of sustainable, inclusive and future-ready higher education systems. At its core, the process begins with students' belief about their own capabilities.

5.1. Limitations

Despite its contributions, several limitations should be acknowledged. First, the review is based on a small number of studies, reflecting the limited empirical work explicitly linking Adlerian principles to self-efficacy in educational settings. This restricts the generalisability of the findings. Second, the included studies varied widely in design, ranging from group counselling interventions to correlational and conceptual papers, which limits the ability to draw firm causal conclusions. Third, because the search strategy relied partly on manual identification of articles, there remains a possibility that relevant studies were not captured due to indexing limitations or terminological variations in how Adlerian concepts are described in the literature. Finally, most interventions were conducted in specific cultural contexts (e.g., Iran, Indonesia, Korea), raising questions about cultural transferability and applicability to broader or more diverse student populations.

Lastly, although the reviewed studies provide valuable insights into Adlerian-informed strategies for enhancing academic self-efficacy, their transferability across cultural contexts remains limited. Most empirical evidence is concentrated within specific cultural and educational settings, which may

influence how Adlerian concepts are experienced and expressed by students. Issues of cultural transferability should be considered when interpreting and applying the findings of this review. While certain Adlerian-based strategies such as encouragement may be relatively vigorous across cultural contexts due to its alignment with universal support processes, other techniques such as early recollections, may require cultural adaptation depending on students' comfort with self-disclosure and their cultural norms.

Therefore, the applicability of these activities may vary across sociocultural settings. As such, the applicability of these findings to diverse higher education contexts should be interpreted with caution. Future study is needed to examine the implementation of Adlerian-based intervention across diverse contexts, including pilot studies with fidelity measures, mixed-methods evaluations, and longitudinal follow-up designs.

5.2. Suggestions for Future Research

Future research should expand the empirical study that incorporates the Adlerian concepts into structured interventions that aims in enhancing self-efficacy. Based on this synthesis, it is recommended that future researcher to consider designing and developing an Adlerian-based intervention that focusing on education or transitional phase among the students as practical tools for counsellors in educational contexts, align with students' need. There is also the necessity for more rigor and diverse research design, such as Randomized Controlled Trial (RCT), longitudinal studies, and mixed methods design in understanding how Adlerian's mechanism functions from time to time.

Future researcher also suggested to explore how Adlerian's principles interact with contemporary counselling approach such as CBT, Solution-Focused Brief Therapy in developing integrative model that were suitable to be used in educational settings. Besides that, more culturally diverse studies are needed to examine how social interest, encouragement, and purposeful living manifest across different educational systems and cultural backgrounds. Future studies also should consider examining specific components such as lifestyle assessment, early recollections, and strength-based strategies, to determine which Adlerian techniques most effectively contribute to changes in students' self-efficacy.

Declaration of AI Use

The authors declare that artificial intelligence-assisted tools were used to support language editing and clarity during manuscript preparation. All conceptual development, data interpretation, synthesis of findings and final content decisions were made by the authors, who take full responsibility for the accuracy and integrity of the manuscript.

References

Abdullah, Z., Ramlan, M. F., Sabran, M. S., & Alsagoff, S. A. (2014). Towards a university branding: The effect of self-efficacy on student development in a major higher education institution. *Jurnal Personalia Pelajar*, 17, 49-61.

Adler, A. (1927). The feeling of inferiority and the striving for recognition. *Proceedings of The-Royal Society of Medicine*, 25-30.

Alipour, F. (2024). The role of university human resources development functions in students' self-efficacy: Empirical evidence from Farhangian University. *The Journal of Theory and Practice in Teachers Education*, 9(3), 97-115. <https://doi.org/10.48310/tpd.2025.17658.1785>

Badejo, A., & Bola, L. S. (2020). Effect of Adlerian therapy on the psychosocial challenges of secondary school adolescents. *European Journal of Educational and Development Psychology*, 8(4), 17-25.

Basileo, L. D., Otto, B., Lyons, M., Vannini, N., & Toth, M. D. (2024). The role of self-efficacy, motivation, and perceived support of students' basic psychological needs in academic achievement. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1385442>

Berezyuk, O., Filts, O., Hrabchak, V., & Sirko, R. (2023). The phenomenon of inferiority and depression. *Proceedings of the Shevchenko Scientific Society. Medical Sciences*, 72(2). <https://doi.org/10.25040/ntsh2023.02.11>

Bilqis, F. (2021). Does Adlerian play therapy increase social, physical, and task appeal more than content mastery services in elementary school students? *Child Education Journal*, 3(1), 1-18. <https://doi.org/10.33086/cej.v3i1.1812>

Bjorklund, P., Daly, A. J., Ambrose, R., & van Es, E. A. (2020). Connections and capacity: An exploration of preservice teachers' sense of belonging, social networks, and self-efficacy in three teacher education programs. *AERA Open*, 6(1). <https://doi.org/10.1177/2332858420901496>

Bourhis, J. (2018). The SAGE encyclopedia of communication research methods. *SAGE Publications*. <https://doi.org/10.4135/9781483381411>

Budiman, R., Purwanto, E., & Murtadho, A. (2020). Cognitive Behavioral Therapy to improve self-efficacy and reduce academic procrastination. *Jurnal Bimbingan Konseling*, 9(3), 145-151. <https://journal.unnes.ac.id/sju/index.php/jubk>

Carnes-Holt, K., Warren, J., Maddox, R. P., Morgan, M., & Zakaria, N. S. (2016). Using bookmarks: An approach to support ethical decision in play therapy. *International Journal of Play Therapy*, 25, 4, 176-185.

Dadandi, I. (2023). Academic self-efficacy, study skills and academic achievement: A serial mediation model. *Research on Education and Psychology*, 7(2), 291-306. <https://doi.org/10.54535/rep.1342022>

Datu, J. A. D., & Yuen, M. (2020). Students' connectedness is linked to higher gratitude and self-efficacy outcomes. *Children and Youth Services Review*, 116. <https://doi.org/10.1016/j.childyouth.2020.105210>

Dias, D., & Sá, M. J. (2014). The impact of the transition to HE: Emotions, feelings and sensations. *European Journal of Education*, 49(2), 291-303. <https://doi.org/10.1111/ejed.12058>

Dortaj, F., & Sohrabi, Z. (2020). The effectiveness of self-encouragement training based on Adler's theory on psychological capital and academic engagement of students. *Quarterly of Clinical Psychology Studies Allameh Tabataba'i University*, 10(37), 91-108.

Döş, İ. (2023). Relationship between happy school, general self-efficacy, academic self-efficacy and life satisfaction. *European Journal of Educational Management*, 6(1), 31-43. <https://doi.org/10.12973/eujem.6.1.31>

Fakhrou, A., & Habib, L. H. (2021). The Relationship between academic self-efficacy and academic achievement in students of the Department of Special Education. *International Journal of Higher Education*, 11(2). <https://doi.org/10.5430/ijhe.v11n2p1>

Farahbakhsh, K., Farshad, M. R., Mirzavand, A., & Riahi, M. (2014). Effectiveness of instructing encouragement, based on Adler's theory, to teachers in interaction with students experiencing academic failure, and its effect on the improvement of students' academic performance. *Online International Journal*, 4(S1), 1987-1993. <http://http://www.cibtech.org/sp.ed/jls/2014/01/jls.htm>

Fatima, Z., & Ashraf, R. (2018). Psychological birth order, self-efficacy and achievement motivation in students. *European Journal of Research in Social Sciences*, 6(6). www.idpublications.org

Firmansyah, A. M. (2025). The relationship between self-efficacy and metacognitive learning strategies in improving student academic achievement. *Educational Praxis Journal*, 1(1), 25-35. <https://doi.org/10.32529/epj.v1i1.3859>

García-Álvarez, D., Cobo-Rendón, R., & Lobos, K. (2024). Character strengths as predictors of general and academic self-efficacy in university students. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1490095>

Gehart, D. R. (2016). *Theory and treatment planning in counseling and psychotherapy* (2nd Ed.). Cengage Learning.

Habsy, B. A., Susanto, A. M., Oetari, A. Y., & Nishafa, D. K. (2025). Adlerian counseling in identifying self-esteem management. *Jurnal Kajian Pendidikan Dan Psikologi*, 3(1). <https://altinriset.com/journal/index.php/jkpp>

Hamann, K., Pilotti, M. A. E., & Wilson, B. M. (2020). Students' self-efficacy, causal attribution habits and test grades. *Education Sciences*, 10(9), 1-14. <https://doi.org/10.3390/educsci10090231>
Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. SAGE.

Jun, Y. (2021). The relationship of academic self-efficacy, goal orientation, and personal goal setting among high school students. *Frontiers in Educational Research*, 4(11). <https://doi.org/10.25236/fer.2021.041109>

Kaplan, V., Düken, M. E., Kaya, R., & Almazan, J. (2023). Investigating the effects of Cognitive-Behavioral-Therapy-Based Psychoeducation Program on university students' automatic thoughts, perceived stress, and self-efficacy levels. *Journal of Research and Health*, 13(2), 87-98. <https://doi.org/10.32598/JRH.13.2.2125.1>

Keller T., & Szakál P. (2021). Not just words! Effects of a light-touch randomized encouragement intervention on students' exam grades, self-efficacy, motivation, and test anxiety. *PLoS One*, 16(9). e0256960 doi: 10.1371/journal.pone.0256960.

Kim, E., & Lee, J. (2025). Effects of Adler's theory-based nudge volunteer program on middle school students. *Psychology in the Schools*, 62(1), 136-148. <https://doi.org/10.1002/pits.23318>

Lee, M.-H., Kim, M.-C., Lee, W.-J., & Leigh An, S.-J. (2025). The impact of group counseling on academic self-efficacy and motivation among Korean university students with academic warnings. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251356079>

Lin, S., Mastrokoulou, S., Longobardi, C., Bozzato, P., Gastaldi, F. G. M., & Berchiatti, M. (2021). Students' transition into higher education: The role of self-efficacy, regulation strategies, and academic achievements. *Higher Education Quarterly*. <https://doi.org/10.1111/hequ.12374>

Manea, V. I., Macavei, T., & Pribeanu, C. (2024). Exploring the influence of academic self-efficacy, engagement, and test anxiety on students' success at exam. *Informatica Economica*, 28(4), 35-44. <https://doi.org/10.24818/issn14531305/28.4.2024.03>

Mathwasa, J., & Sibanda, L. (2020). Enhancing students' self-efficacy: Implication for high school guidance and counselling educators. In S. G. Taukeni, *Counseling and Therapy*. IntechOpen.

Meng, Q., & Zhang, Q. (2023). The influence of academic self-efficacy on university students' academic performance: The mediating effect of academic engagement. *Sustainability*, 15(7), 57-67. <https://doi.org/10.3390/su15075767>

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey Bass.

Michael, F. L., Sumilan, H., Bandar, N. F. A., Hamidi, H., Jonathan, V., & Md Nor, N. N. (2020). Sustainable development concept awareness among students in higher education: A preliminary study. *Journal of Sustainability Science and Management*, 15(7), 113-122. <https://doi.org/10.46754/jssm.2020.10.011>

Pedditz, M. L. (2024). School satisfaction and self-efficacy in adolescents and intention to drop out of school. *International Journal of Environmental Research and Public Health*, 21(1). <https://doi.org/10.3390/ijerph21010111>

Quality Assurance Agency Scotland. (2023). Transition skills and strategies: Self-efficacy. QAA Scotland. <https://www.qaa.ac.uk>

Püttmann, V., & Gessner, S. (2023). Young people's perceptions of transitioning from school to higher education: A cross-national perspective. *International Journal of Education and Literacy Studies*, 11(1), 53-65. <https://doi.org/10.7575/aiac.ijels.v.11n.1p.53>

Rismawan, K. S., & Gading, K. (2021). The effectiveness of Cognitive Behavior Group Counseling to improve career decision making self-efficacy of senior high school students. *Advances in Social Science, Education And Humanities Research*, 540.

Safara, M., Rezaei Nasab, A., & Gooshki, H. S. (2020). The effectiveness of self-encouragement training on reduce aggression and increase the academic self-efficacy of high school students. *Pakistan Journal of Medical and Health Sciences*, 14(3).

Sides, J. D., & Cuevas, J. A. (2020). Effect of goal setting for motivation, self-efficacy, and performance in elementary mathematics. *International Journal of Instruction*, 13(4), 1-16. <https://doi.org/10.29333/iji.2020.1341a>

Stipanovic, N., Stringfield, S., & Witherell, E. (2017). The influence of a career pathways model and career counseling on students' career and academic self-efficacy. *Peabody Journal of Education*, 92(2), 209-221. <https://doi.org/10.1080/0161956X.2017.1302217>

Subarimaniyam, N., Zakaria, N. S., & Wan Jaafar, W. M. (2020). Multicultural competency, spirituality and self-efficacy in dealing with legal and ethical issues. *Pertanika Journal for Social Sciences and Humanities*, 28, 2, 1371-1386.

Sukhera, J. (2022). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education*, 14(4), 414-417. <https://doi.org/10.4300/JGME-D-22-00480.1>

Sundah, A. J. A. (2019). Adlerian group counselling to improve student's self-efficacy. *International Journal of Innovation, Creativity and Change. Www.Ijicc.Net*, 5(5), 136-147. www.ijicc.net

Svartdal, F., Sæle, R. G., Dahl, T. I., Nemtcen, E., & Gamst-Klaussen, T. (2022). Study habits and procrastination: The role of academic self-efficacy. *Scandinavian Journal of Educational Research*, 66(7), 1141-1160. <https://doi.org/10.1080/00313831.2021.1959393>

Sweeney, T. J. (2009). *Adlerian counselling and psychotherapy: A practitioner's approach* (5th Ed.). Routledge.

Thompson, M., Pawson, C., & Evans, B. (2021). Navigating entry into higher education: The transition to independent learning and living. *Journal of Further and Higher Education*, 45(10), 1398-1410. <https://doi.org/10.1080/0309877X.2021.1933400>

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, 14, 207-222.

Türk, F., & Katmer, A. N. (2019). A study on the effectiveness of coping with test anxiety program based on cognitive-behavioral approach. *International Journal of Evaluation and Research in Education*, 8(4), 666-675. <https://doi.org/10.11591/ijere.v8i4.20316>

Vachova, L., Sedlakova, E., & Kvintova, J. (2023). Academic self-efficacy as a precondition for critical thinking in university students. *Pegem Egitim ve Ogretim Dergisi*, 13(2), 328-334. <https://doi.org/10.47750/pegegog.13.02.36>

Wahid, A., & Neviyarni, N. (2024). Guidance and counseling efforts to improve students' academic self-efficacy: A systematic literature review. *Counsnesia Indonesian Journal Of Guidance and Counseling*, 5(2), 165-171. <https://doi.org/10.36728/cijgc.v5i2.4188>

Warren, J., Zavaschi, G., Covello, C., & Zakaria, N. S. (2012). The use of bookmarks in teaching counseling ethics. *Journal of Creativity in Mental Health*, 7, 2, 187-2012.

Wong, Y. J. (2015). The psychology of encouragement: Theory, research, and applications. *The Counseling Psychologist*, 43(2), 178-216. <https://doi.org/10.1177/0011000014545091>

Yang, Y., Sriwisathiyakun, K., & Li, J. (2025). Advancing higher education students' self-efficacy and achievement through self-directed learning integrated with the BOPPPS model. *Cogent Education*, 12(1), 2554320. <https://doi.org/10.1080/2331186X.2025.2554320>

Yu, W., Zhou, S., & Zhou, Y. (2024). Intervention on Mathematics self-efficacy: Solution-Focused Brief Therapy. *Psychology Research and Behavior Management*, 17, 129-145. <https://doi.org/10.2147/PRBM.S432569>

Zakaria, N. S., Saripan, M. I., Subarimaniyam, N., & Ismail, A. (2020). Assessing Ethoshunt™ as a gamification-based mobile application in ethics education: Pilot mixed-methods study. *Journal of Medical Internet Research-Serious Games Journal*, 8, 3: e18247, doi: 10.2196/18247.

Zakaria, N. S., Subarimaniyam, N., Wan Jaafar, W. M., Mohd Ayub, A. F., & Saripan, M. I. (2022). Conceptualization and initial measurement of counseling ethics competency: The influence of spiritual and self-efficacy. *European Journal of Training and Development*, 46, 3-4, 285-301(17), doi: 10.1108/EJTD-02-2020-0016.

Zakaria, N. S., & Warren, J. (2016). Counseling ethics education: Teaching and learning development reformation in I. H. Amzat and B. Yusuf (Eds.), *Fast Forwarding Higher Education Institutions for Global Challenges: Perspectives and Approaches*, 83-96, Springer, doi: 10.1007/978-981-287-603-4.

Zakaria, N. S., Warren, J., & Bakar, A. R. (2017). Counseling ethics education for enhanced professional identity and development: Guidance and counseling teachers lifelong learning acquisition empowered in I. H. Amzat and N. P. Valdez (Eds.), *Teacher Empowerment for Professional Development and Practice: Perspectives Across Borders*, 153-166. Springer, doi: 10.1007/978-981-10-4151-8.

Žalėnienė, I., & Pereira, P. (2021). Higher education for sustainability: A global perspective. *Geography and Sustainability*, 2(2), 99-106. <https://doi.org/10.1016/j.geosus.2021.05.001>

Zayed, A. M. (2024). Academic self-efficacy and its relationship to academic competitiveness, academic procrastination, and cognitive flexibility among undergraduate students. *International Journal of Higher Education*, 13(1), 57. <https://doi.org/10.5430/ijhe.v13n1p57>

Zeidi, I. M., Divsalar, S., Morshedi, H., & Alizadeh, H. (2020). The effectiveness of group cognitive-behavioral therapy on general self-efficacy, self-control, and internet addiction prevalence among medical university students. *Social Health and Behavior*, 3(3), 93-102. https://doi.org/10.4103/SHB.SHB_20_20

Appendix A: Full search string for each database

Database	Search string
Scopus	(Adler OR "Adlerian counseling" OR "Adlerian counselling") AND ("self-efficacy" OR "academic self-efficacy")

	AND (education OR school OR university OR students)
SAGE	(Adlerian OR Adler) AND ("self-efficacy" OR "academic self-efficacy") AND (education OR school OR university OR students)
ERIC	("Adlerian counseling" OR "Adlerian counselling" OR Adler) AND ("self efficacy" OR "academic self efficacy") AND (education OR school OR university OR students)
Google Scholar	"Adlerian counselling" OR "Adlerian counseling" "self-efficacy" education

Learning Online Through Community of Inquiry and Culturally Responsive Teaching (COI-CRT) Framework: Insights from Chinese Students Studying in Malaysian Higher Education

ABSTRACT

Authors: Wang Yu & Nur Aira Abdrahim

Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:

nuraira@upm.edu.my

Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026: |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

The internationalisation of higher education has led to a growing number of Chinese students pursuing academic opportunities abroad, including in Malaysia, one of the region's major education hubs. At the same time, online learning has emerged as a mainstream form of instructional delivery, requiring international students to adapt not only to new academic environment, but also studying within an online learning system. This study examines how Chinese students in Malaysian higher education experienced online learning through the interpretive lens of the Community of Inquiry (CoI) framework, focusing on teaching, cognitive, and social presences with the integration of Culturally Responsive Teaching (CRT) framework. Using a qualitative design, interviews were conducted with a total of ten Chinese international students from two public universities. Data were analysed inductively, with the CoI-CRT framework serving as a sensitising concept to guide interpretation. Findings show that students navigated online learning through a personal check and balance process that involved monitoring their understanding, managing flexibility, and compensating for the absence of in-person structure. Teaching presence strongly shaped their experiences; with clear organisation, timely responses, culturally relevant examples, and technological guidance were key enablers of productive engagement. Cognitive presence was enhanced when instructors used simulations, case-based examples, and interactive activities that encouraged deeper thinking. However, cognitive engagement was easily disrupted by technical issues such as lagged lectures and unstable internet connections. Social presence emerged as the most challenging dimension. Students reported reduced interaction, limited nonverbal cues, and uncertainty in multilingual discussions, which contributed to feelings of disconnection. Different cultural norms can further heighten fear and hesitancy among the learners to participate in online settings. The study underscores the need for culturally responsive teaching (CRT) practices within online environments, where intentional design and facilitation can foster inclusion, reduce uncertainty, provide safe space and support more meaningful engagement for diverse international learners.

Keywords: Online learning, international students, Community of Inquiry, culturally responsive teaching.

1. Introduction

Tertiary education is currently going through a rapid globalization that is driven by academic and economic imperatives to meet internationalisation standard. Internationalisation in this context is defined as the intentional process of integrating international, intercultural, and global dimensions into the purpose, functions, and delivery of post-secondary education (Knight, 2004). This definition underscores the fact that internationalisation in higher education does not happen at the periphery—it is a strategic, planned, and systemic transformation across teaching, research, community engagement, and institutional operations. Internalisation has also become a strategic priority in the current knowledge economy, thus reshaping how universities design and deliver their programmes in order to attract students around the world who are looking to pursue their education abroad.

This global push towards internationalisation has encouraged many nations to strengthen their higher education systems to compete internationally, Malaysia included. Therefore, many universities have increasingly integrate online and blended learning approaches into their curricula, reflecting a broader national shift toward digitalised teaching and learning environments (Kathirveloo, 2024). As one of the prominent destination for international students, Malaysia has consistently ranked among the top choices for China students due to its affordability, cultural proximity, and quality of higher education. Malaysia's position as a major educational hub in Southeast Asia has received over 130,000 overseas students annually, with Chinese students constituting about 41% of this population (Ministry of Higher Education, 2025).

Chinese students' motivations for studying abroad are shaped by a combination of academic aspirations, social expectations, and structural pressures within China's competitive education system (Liu & Morgan, 2017). They highlight that many Chinese students pursue overseas education to access higher-quality programmes, international learning environments, and obtain degrees that are perceived to have stronger global recognition. Their decision-making is also influenced by family expectations, where these Chinese parents view foreign education as an investment in social mobility and future employability of their children. In addition, the limited availability of places in top-tier Chinese universities also serves as a strong push factor, driving students to seek opportunities abroad.

Despite their strong motivation to pursue studies abroad, Chinese international students often encounter significant challenges when adjusting to overseas higher education environments. Research shows that they frequently struggle with language barriers, especially in English-medium classrooms where unfamiliar academic vocabulary may limit active participation in academic discussions (Zhou & Todman, 2009). Many also find it difficult to adapt to more interactive, student-centred pedagogies that contrast with the teacher-directed approaches common in China, leading to perceived gaps in their participation in classroom engagement and critical discussion (Heng, 2018). Social integration is another hurdle, as cultural differences, communication norms, and feelings of isolation can hinder meaningful connections with peers and instructors. When in online learning contexts, these challenges are further amplified by reduced opportunities for informal interaction, uncertainties around self-directed learning, and the need to navigate new digital platforms independently (Castulo, 2025; Tu, 2010).

These challenges align with broader patterns observed among international students, who must navigate psychological, sociocultural, and academic transitions as part of their adjustment process (Rahim, 2021). More importantly, Rahim (2021) found that academic adjustment is viewed a central priority for international students, influencing how they navigate and make meaning of their overall adjustment process so that they can function academically. The emphasis on academic functioning suggests that, for many international students, adjustment is not only focuses on cultural or social adjustment but deeply connected to their ability to learn, participate, and meet academic demands. Such priorities become even more pronounced in online learning contexts, where students must adapt to new modes of instruction, digital platforms, and independent learning expectations.

In this study, the interest to focus on the Chinese students' experiences in learning online is motivated by the rapid internationalisation and digitalisation in Malaysian higher education practices. As China remains as one of the largest international student groups globally (OECD, 2025), understanding their learning experiences is crucial in enhancing the quality and inclusivity of Malaysian higher education to remain globally competitive. Currently, there is an underexplored area where the intersection of cultural factors of these Chinese students rarely being examined together with online instructional design in Malaysia higher education context. Therefore, this study wishes to frame their online learning experience from the lens of Community of Inquiry (CoI) framework (Garrison, 2017) and Culturally Responsive Teaching (CRT) (Gay, 2010).

The CoI framework bring forth three intersecting elements, namely social presence, cognitive presence, and teaching presence, to help illuminate the context and complexities of online learning as experienced by these Chinese students. Meanwhile, culturally responsive teaching (CRT) emphasises recognising learners' cultural backgrounds, validating their ways of knowing, and designing instruction that supports diverse participation and learning needs (Gay, 2010). In online settings, these principles become especially important, as students' sociocultural adjustment and academic engagement are shaped by differences in communication norms, learning expectations, and familiarity with interactive digital environments.

Insights from this study, interpreted through the Community of Inquiry framework, point to specific areas—such as impactful teaching presence, culturally sensitive facilitation to strengthen social presence, and designing relevance instructional design to deepen cognitive presence—when bridged together with the CRT framework. Understanding how Chinese international students experience online learning is critical for strengthening culturally responsive teaching practices in Malaysian higher education. These insights can inform more inclusive online teaching approaches that not only address common challenges faced by Chinese international students but also enhance the broader institutional goal of fostering culturally responsive and supportive learning environments in higher education.

2. Literature Review

2.1. Learning Online in Cross-Cultural Contexts

Online learning is a form of education where the majority of content delivery, interaction, and assessment occurs through web-based systems or internet-based platforms, allowing flexible access to learning experiences independent of time or location (Moore, Dickson-Deane & Galyen, 2011). In the post COVID-19 era, online learning has now become an integrated approach used by many higher education instructors in their curriculum delivery to provide students with access to different knowledge, resources and ways of learning. It is no longer just a method of delivery aimed for those who are learning at distance, but as part of innovative approach to learning, reflecting global shifts toward digitalisation and technology-supported instruction.

In Malaysia, similar development are evident as higher education institutions expand their use of technology-enhanced learning in line with national priorities for digital transformation. The integration of online platforms, learning management system, and digital assessment systems has become common across public and private universities, shaping everyday teaching practices and student learning experiences (Muniisvaran, Jose, & Anusia, 2025). These platforms support a range of academic activities—from accessing materials and submitting assignments to participating in discussions and collaborative tasks—reinforcing the role of online environments as routine components in the curriculum delivery. Therefore, online participation is now expected among both local and international students to fulfil their academic expectations, engage with instructors, and collaborate with peers.

Nevertheless, for overseas students, learning online added another layer of unfamiliarity arising from the cross-cultural nature of their academic experience abroad. Adapting to online learning requires understanding new technological platforms, communication norms, and participation expectations, but

for students studying outside their home country, these demands intersect with the broader process of navigating different cultural, linguistic, and new educational environments (Zhou & Todman, 2009; Heng, 2018). Cultural norms around classroom interaction, questioning, collaboration, and teacher–student relationships vary across educational systems, and these differences can become even more pronounced in online settings where cues are reduced and expectations are less explicit. Studies note that international students must simultaneously negotiate technological adjustment and sociocultural adaptation, making their online learning experiences more complex (Castulo, 2025).

Research consistently shows that Chinese students may encounter additional challenges when adjusting to online learning, including differences in communication norms, hesitation toward open classroom discussion, and a preference for structured teacher guidance—patterns rooted in their prior educational experiences and cultural expectations of learning (Heng, 2018). These cultural orientations can influence how they interpret teaching presence, participate in peer interactions, and construct meaning in online learning environments. A core element affecting Chinese students' adaptation is the cultural diversity between their home and host educational backgrounds. Hofstede's (1980) seminal work on cultural dimensions recognized collectivism as a key value in Chinese culture, forming students' study actions, including dependence on group work and respect for authority figures (Triandis, 1995). By contrast, as influenced by the Western pedagogical models, Malaysia's education approaches tend to integrate more student-centred learning approaches with active involvement expected from the learners. These contrasting values can result in the feeling of dissonance among the China international students during the online learning process.

Studies further indicate that Chinese international students often report difficulties establishing social presence and navigating unfamiliar learning platforms, especially when studying abroad where both the academic culture and digital expectations differ from those in China (Castulo, 2025). Language proficiency adds another layer of complexity to their learning. Although a lot of Chinese students are skilled at written English, the conversational capabilities can be hampered by language anxiety. This can result in decreased engagement and missed chance for collaborative or student-led learning. According to studies by Cao, Zhu, and Meng (2021), language difficulties worsen pressure among Chinese students, so that they are less likely to seek assistance or ask any questions in online environments.

The learning platform is also important in forming students' online study experiences. As Venkatesh et al. (2003) believed, perceived ease of application and usefulness are important in deciding users' acceptance of technology. For Chinese students used to platforms such as WeChat and QQ, adapting to using different communication platforms in Malaysia such as emails or forums, may be perceived as unfamiliar. The absence of proper training and instruction often results in frustration, decreasing the effect of online study contexts (Ng & Fang, 2023). Social separation is also an ongoing issue in online learning, especially for overseas students who are isolated from their familiar surroundings. With regard to Chinese students, limited spontaneous communication that can happen during online learning can limit their opportunities for peer interactions, thus making their online learning process less effective (Tu, 2010).

Educational organizations in Malaysia have included some measures to build a more inclusive online learning setting. Nevertheless, Heng (2018) discover that the success of these involvements often relies on their capacity of conforming to the cultural expectations, linguistic abilities and technical capabilities of these Chinese students. Gaining a clearer understanding of these factors is essential for improving their online learning experiences, particularly as it highlights the need for teaching approaches that are sensitive to students' cultural background and linguistic ability. In this sense, incorporating elements of culturally responsive teaching can help ensure that online learning supports and includes Chinese learners more effectively.

2.2. Culturally Responsive Teaching (CRT) in Online Context

Modern internationalisation emphasises educational quality, inclusivity, and student support, not merely recruitment. Scholars argue that internationalisation is not only about increasing student mobility alone, but also about creating inclusive learning environments where culturally diverse students can participate meaningfully and succeed academically (Jones, 2013; Leask, 2015). Culturally responsive pedagogy plays a central role in this process by drawing on students' cultural knowledge and experiences, acknowledging differences in communication norms, and designing learning activities that promote equitable participation (Gay, 2010; Ladson-Billings, 1995). Such approaches are especially important in higher education contexts that enrol large numbers of international students, as these learners often navigate unfamiliar academic conventions while simultaneously managing sociocultural adjustment demands.

Culturally responsive teaching (CRT) is grounded in the recognition that students' cultural backgrounds, identities, and ways of knowing significantly shape how they engage with learning. Gay (2010) defines CRT as the use of cultural knowledge, prior experiences, frames of reference, and performance styles of ethnically diverse students to make learning encounters more relevant and effective for them, emphasising that teaching should build on rather than ignore learners' cultural resources. In recent years, CRT has increasingly been discussed in relation to higher education and international students. Studies highlight that culturally responsive pedagogy can help create more inclusive learning environments.

Furthermore, recent scholarship has begun to emphasise the value of using CRT to analyse online learning in internationalised or multicultural higher education systems. Mahani (2023) argues that applying culturally relevant pedagogy to online learning requires intentional design of course structures, activities, and communication practices that support diverse students' engagement and sense of inclusion in digital spaces. Similarly, Khoo and Huo (2022) demonstrate that culturally responsive online support programmes can enhance international students' academic language development and confidence by aligning instructional approaches with students' cultural and linguistic needs. These studies collectively suggest that CRT can inform the design of online courses that are not only technically accessible but also responsive to varied cultural expectations of teaching, interaction, and knowledge construction.

2.3. The Community of Inquiry (CoI) Framework as a Lens for Online Learning

Within this broader literature, the principles of CRT intersect meaningfully with the Community of Inquiry (CoI) framework. Practices that enhance teaching presence—such as making expectations explicit, providing culturally sensitive feedback, and structuring clear guidance—can be viewed as culturally responsive when they take into account students' cultural backgrounds and prior learning experiences. Likewise, efforts to foster social presence by validating diverse voices, encouraging equitable participation, and designing collaborative tasks that respect different communication styles align with CRT's emphasis on inclusive and affirming learning communities. As online learning becomes a standard mode of delivery in Malaysian higher education, integrating CRT perspectives into the analysis of teaching, social, and cognitive presence offers a useful pathway to better understand and support the experiences of culturally diverse student groups, including Chinese international students.

The Community of Inquiry (CoI) framework is one of the most widely adopted models for examining and designing meaningful online learning experiences in higher education. First introduced by Garrison, Anderson, and Archer (2000), the framework conceptualises online learning as a collaborative process in which knowledge is constructed through the dynamic interaction of three core elements: teaching presence, social presence, and cognitive presence. These presences work together to shape the quality of learners' engagement, the coherence of instructional design, and the depth of understanding that develops in virtual learning environments (Garrison, 2017). The CoI framework therefore offers a structured lens for analysing how students experience online learning, particularly in contexts where interaction, communication, and guidance occur primarily through digital platforms.

Teaching presence refers to the design, facilitation, and direction of learning activities, and is central to establishing clarity, structure, and guidance in online courses. Research shows that strong teaching presence supports learner confidence, reduces uncertainty, and anchors students' navigation of course expectations—features especially important in environments where immediate, in-person cues are absent (Garrison & Arbaugh, 2007). Social presence, defined as learners' ability to project themselves as real and authentic individuals in online environments, enables the development of trust, connection, and expressive communication among participants. Studies indicate that social presence contributes to students' satisfaction and learning outcomes (Richardson et al., 2017). Lastly, cognitive presence captures the extent to which learners construct meaning through reflection and discourse. It represents the intellectual dimension of the model, emphasising inquiry, exploration, and sustained engagement with ideas.

The CoI framework is particularly relevant for examining online learning within culturally diverse settings because it highlights the relational, instructional, and cognitive conditions needed for effective participation. In cross-cultural learning environments, differences in norms related to communication, authority, peer interaction, and classroom participation may shape how teaching, social, and cognitive presences are perceived and enacted. For international students—who must simultaneously adapt to unfamiliar academic expectations and digital modes of learning—the CoI lens helps illuminate where misalignments or challenges arise, such as limited social presence due to differing communication styles, or uncertainties about teaching presence when expectations are not made explicit. As Garrison (2017) notes, the CoI model serves as a lens that illuminates the complexities of online learning, making it especially suitable for understanding the layered realities of learners navigating new cultural and academic terrains.

The expansion of online learning has prompted scholars to consider how CoI elements can be translated into this learning environment. For instance, Arbaugh's (2013) study looks at how specific academic discipline of a course influences the effectiveness of teaching strategies across 51 online MBA courses. The research demonstrates that common, standardized prescriptions for designing online courses must have considered the context of its academic disciplines. Although not specifically examined in multicultural context, Arbaugh's study offers insights on how CoI elements are applied in multidisciplinary context. Likewise, a study on blended learning design for intercultural communication courses in Chinese universities demonstrated that applying the Community of Inquiry (CoI) framework can enhance students' intercultural competence, including their knowledge, attitudes, and awareness, while also improving engagement and learning behaviours (Wei, 2024).

These insights align closely with the goals of culturally responsive teaching, reinforcing the idea that online learning environments must accommodate diverse academic identities, communication norms, and learning expectations. In this study, the integration between CoI and CRT provides a robust theoretical base for examining the experiences of Chinese international students and identifying practices that enhance their engagement and learning success within online context. The intersection of these two theories forming the conceptual integration for this study is provided in Figure 1.

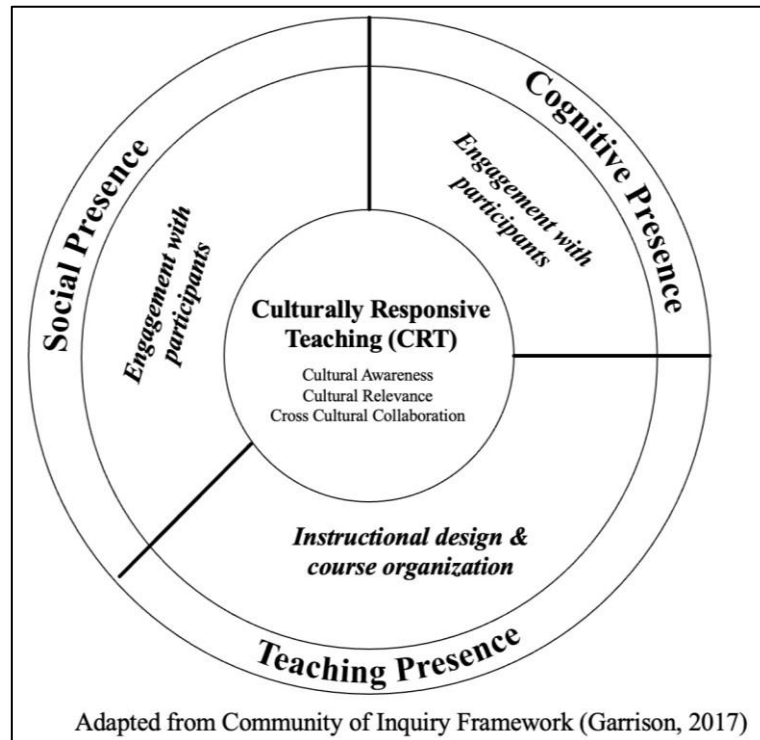


Figure 1. Conceptual Integration of Community of Inquiry (CoI) with Culturally Responsive Teaching (CRT) Framework

As illustrated in the conceptual figure, CRT is positioned at the centre of the model as an integrative layer that informs and shapes how each CoI presence is enacted. Meanwhile, the outer structure represents the three core elements of the CoI framework—teaching presence, social presence, and cognitive presence the three key areas that must be considered in designing effective and meaningful online learning experiences (Garrison, 2017). The circular design of the figure indicates the dynamic and interdependent relationship between the three presences, with CRT embedded across all dimensions, rather than functioning as a separate construct. This integration highlights that effective online learning for international students is not solely dependent on instructional design alone, but also on how teaching, social and cognitive presence are considered as contextually and culturally relevance.

3. Methodology

3.1. Research Design

This study adopted a qualitative research design to explore how Chinese international students experience online learning in Malaysian higher education. Qualitative research is fundamentally concerned with understanding how individuals interpret and make meaning of their experiences within specific social, cultural, and contextual conditions. Rather than seeking to measure variables or generalise findings to large populations, qualitative inquiry prioritises depth, richness, and the exploration of phenomena as lived and perceived by participants (Merriam & Tisdell, 2016). Given that online learning is embedded within broader sociocultural and academic adjustment contexts for international students, a qualitative design enables a nuanced exploration of how these intersecting factors shaped the Chinese learners' experienced when learning online through CoI-CRT framework. Therefore, a qualitative approach is well-suited for investigating complex, context-dependent phenomena such as teaching presence, social presence, and cognitive presence, which are best understood through meaning-making processes and experiences as reported by these learners. By foregrounding students' voices, the study prioritises depth over breadth and seeks to understand the complexities of learning online from the perspective of those who experience it directly. This approach

aligns with calls in the literature for more interpretive, culturally aware investigations into international students' online learning experiences, particularly within Asian higher education contexts.

3.2. Location and Sampling

This research recruited participants from two public universities in Malaysia, targeting universities with large recruitment of China international students and have been identified as educational institutions that integrate online learning as part of their instructional design in curriculum delivery. Sampling from two universities allowed the study to capture broader online learning experiences that may arise from differences in programme structures, online teaching practices, and learning management systems used across different learning institutions, while still situating the research within the shared academic context of Malaysian public higher education.

Within these two public universities, selection of participants was recruited based on purposive sampling, where researchers select samples where the study can learn from the most (Merriam & Tisdell, 2016). Furthermore, criterion sampling was used to identify study participants that meet some predetermined criterion of importance (Patton, 2015), ensuring that all samples selected are directly relevant to meet the purpose of the study. The inclusion criteria include i) International students from People's Republic of China (PRC); ii) Enrolled as full-time students at a Malaysian public university as undergraduate or postgraduate students only (excluding mobility or exchange students); iii) Have participated and learned in online or blended learning for at least two semesters in courses offered at their university. During the open recruitment process, a total of 55 emails and text invitations have been sent out, with twenty-three favourable responses have been received initially. Six participants opted out due to language requirements, two was excluded due to being exchange students, while the remaining was not included in the interview process due to the unavailability of time. In total, 10 participants were recruited as part of this study's sample, consisting of both undergraduate and graduate international students from China, who expressed willingness to participate in this study and met the study inclusion criteria as listed. Please see Table 1 below for the participant's demographic profiles.

Table 1. Participants Demographic Profiles

Participant ID	Gender	Age	Program Major	Program Level
P01	Male	22	Business	Undergraduate
P02	Male	24	Engineering	Postgraduate
P03	Female	23	Education	Postgraduate
P04	Female	26	Law	Postgraduate
P05	Male	25	Computer Science	Postgraduate
P06	Female	21	Business	Undergraduate
P07	Female	23	Economics	Undergraduate
P08	Female	22	Economics	Undergraduate
P09	Female	24	Psychology	Postgraduate
P10	Male	23	Education	Postgraduate

3.3. Data Collection Methods

In this study, interviews served as the primary method of data collection. Semi-structured interviews were chosen specifically for the flexibility and depth allow by this data collection method, enabling the researchers to guide the conversation with predetermined questions while also allowing the participants to express their individual's experiences in their own words (Merriam & Tisdell, 2016).

Interviews were conducted individually to allow participants the privacy needed to share personal reflections about their experiences navigating and learning online. The interviews were conducted by the first researcher with the presence of the senior researcher (as observer), therefore it was primarily conducted in basic English for consistency. Mandarin was only used in case the participants did not understand the question asked, but all responses were requested in English. Each interview lasted between 45 to 60 minutes and was conducted either face-to-face or via Zoom, depending on participants'

preferences and availability. Conducting interviews online was especially useful when participants' schedules were constrained or when they preferred to demonstrate their learning platforms directly using screen sharing function.

Furthermore, an interview protocol was developed to provide balance between ensuring coverage of key areas of the study's interest while also allowing participant-led conversation to occur. Rather than imposing a rigid structure on the protocol, the interview questions were intentionally designed to be general and conversational, while avoiding jargon and abstract academic terminology. The protocol was developed to follow a three-stage flow that moves from general to specific, with the purpose of supporting rapport-building and allowing participants to ease from the general conversation to deeper reflection (Merriam & Tisdell, 2016).

In the first part, the interviews opened with introductory questions about participants' background, academic programme, and general impressions of online learning in Malaysia. In the second stage, participants were asked to describe their online learning experiences in more details, focusing on the factors contributing to their positive online learning experiences and the challenges they encountered. To ensure sufficient coverage of the interview guide meeting the objectives of the study, questions were developed as guided by the corresponding research questions (Please see Table 2).

Table 2. Samples of interview questions

Corresponding RQ	Selected Samples of Interview Questions and Probes
RQ1: How do Chinese students describe their online learning experiences at public university in Malaysia?	1. How do you feel about studying online as an international student in Malaysia? 2. In what ways is learning online similar to or different from your previous learning experiences? 3. Can you give some examples of what you just described?
RQ2: What are the factors that contribute to the Chinese students' positive online learning experiences?	1. What aspects of online learning have helped you learn effectively? 2. What kinds of activities or tasks made online classes more engaging or meaningful for you? 3. Can you describe any teaching practices or instructor behaviours that supported your learning online?
RQ3: What are the challenges that Chinese students faced in online learning experiences?	1. What difficulties did you encounter while learning online? 2. How did studying online as a Chinese international student challenge your learning process? 3. What are your suggestions to deal with these challenges?

Lastly, the third stage involved an online learning platform walkthrough, where participants were invited to demonstrate how they navigated their Learning Management System (LMS) or course platforms as part of their online learning experiences. This stage was included for gaining contextual insight that can reveal aspects of their online experience that participants may struggle to describe verbally, which tends to occur among international students especially due to language and terminology barriers.

Data collection continued until thematic saturation was reached, defined as the point at which no substantially new themes or codes emerged from additional interviews (Saunders et al, 2018). Saturation was assessed through ongoing comparison of data during the analysis process, guided by several criteria, including increasing repetition and redundancy in participants' responses and the absence of new codes in the coding framework (Guest et al., 2006; Strauss & Corbin, 1998). In this study, repetition began to occur after the eighth participant, with no new codes identified in two consecutive interviews thereafter. Subsequent interviews served to confirm and deepen existing themes rather than generate new insights, indicating that sufficient depth and coverage of the phenomenon had been achieved. The process of data management and data analysis followed as described in the next sections.

3.4. Data management and trustworthiness

This study adheres to the ethical consideration as outlined by the University Ethics Committee for Research Involving Human Subjects (JKEUPM), aligned with international ethical standard guidelines as articulated in the Belmont Report (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979) and the British Educational Research Association Ethical Guidelines for Educational Research (British Educational Research Association, 2018). Therefore, all safeguard measures have been applied in the study following internationally recognised ethical principles governing qualitative research in data management and for trustworthiness purposes.

First, all participants provided informed consent after receiving clear information about the study's purpose, procedures, and their rights as voluntary participants. Additionally, pseudonyms were assigned to all participants, and any identifiable details—such as programme, year of study, or references to specific instructors—were removed or generalised in the data transcription and during reporting.

All data were transcribed and managed through NVivo software (Version 12) in order to facilitate a systematic data management and analysis process as described below. All data properties inside NVivo are password-protected to safeguard access only to the researchers. In accordance with recognised research data management practices, these data will be retained for five years for audit trail purpose.

Several strategies were incorporated to strengthen the study's trustworthiness. All participants were invited to take part in the member check process to evaluate the accuracy of their narratives. Additionally, a peer debriefing process also took place between the first researcher with the senior researcher who has prior experience in conducting qualitative research and online learning studies, but not directly involved in the study. Given the first researcher's shared cultural background with the participants, this process was used as a mechanism to minimise researcher bias to ensure credibility and reliability of the data collected. Particular attention was given to the flow of interactions during the interview, clarity of questions asked as well as identifying any instances where questions may have been unintentionally leading or framed in ways that could influence participants' responses.

The debriefing process also extended during to the analysis process as the primary coding and analysis process were conducted by the first researcher who is also a Chinese international student studying in Malaysia. Due to the primary coder's positionality in connection with the topic under study, conscious efforts were made to keep the researcher bias in check by practising reflexivity and keeping audit trail throughout the research process. During these sessions, emerging codes, theme definitions, and illustrative excerpts were reviewed and critically discussed. This process helped challenge assumptions, refine theme boundaries, and ensure that interpretations were grounded in the data rather than researcher bias. The next section outlines the details of the data analysis that took place in this study.

3.5. Data Analysis

In this study, the data analysis was approach while recognising the importance of subjective and open interpretation in qualitative inquiry. Therefore, to remain open in the interpretation of findings, the data analysis was conducted using inductive thematic analysis to elicit themes grounded in the data that are being driven by the participants' experiences, using the steps recommended by Braun and Clarke (2006).

To ensure theoretical rigor in the analysis, the CoI framework was applied as a sensitising concept to provide an interpretive frame that orients the researchers to certain features of the data while preserving its openness (Bowen, 2006). Therefore, in the initial phase, open coding analysis was conducted to provide holistic understanding of each participant's narrative, staying close to their language and expressions. This allowed early patterns, meanings, and tensions to surface organically from the data, ensuring that the analysis remained grounded in the participants' words. All these processes were done recursively by revisiting the coding logs until the corresponding themes were satisfactorily refined and accurately reflective of the participants' experiences. Only after these inductive themes were established, the CoI-CRT framework was brought in as a sensitising concept to help interpret how

elements of teaching presence, social presence, and cognitive presence were reflected in the emerged themes. As noted by Bowen (2006), sensitising concept enhances the trustworthiness of a qualitative research as it provides conceptual grounding to the interpretation while preserving the analytical openness of the findings. The overall steps conducted in the study's analysis are outlined under Table 3.

Table 3. Steps of Data Analysis

Analytic Step	Description of Analytic Activity	Source of References
Familiarisation	Transcribing and repeated reading of transcripts to gain an overall sense of participants' experiences	Braun & Clarke (2006)
Generating initial codes	Line-by-line coding conducted inductively, staying close to participants' own words	Braun & Clarke (2006)
Searching for themes based on RQs	Re-focuses the analysis at the broader level of themes and producing preliminary thematic mapping of findings	Braun & Clarke (2006)
Reviewing the themes through sensitising process	Introduction of CoI and CRT as interpretive lenses to refine themes and examine theoretical linkage	Braun & Clarke (2006) Bowen (2006)
Defining and naming finalized themes	Themes clarified and articulated with conceptual depth while retaining inductive grounding	Braun & Clarke (2006)
Producing the report	Findings presented and discussed in relation to CoI and CRT literature	Braun & Clarke (2006) Garrison (2017) Gay (2010)

To maintain consistency during data analysis process, a coding scheme was developed to guide the sensitizing process. Drawing upon the CoI–CRT framework, these sensitising concepts provide an interpretive lens rather than fixed categories or themes for the findings. This process ensures the theoretical rigor was applied consistently during the analytical process. The following Table 4 outlined the definitions of key constructs, guided analytical interpretations, and a corresponding key indicators that informed how these key constructs were recognized in the data and used as basis to realign the findings with the CoI-CRT theoretical framework.

Table 4. Key constructs and guided analytical interpretations of CoI-CRT model during data sensitising process

Theoretical Framework	Key Construct	Guided Analytical Interpretations (adapted for online learning context)	Key Indicators
CoI	Teaching presence	How instructors design, organise, and facilitate online learning experiences, including the structuring of online content, guidance of learning activities, and maintenance of instructional clarity within online learning environment.	1. Clear organisation and explicit instructions for online tasks and assessments 2. Instructor responsiveness to questions during online and offline session 3. Timely and constructive feedback provided through digital platforms like LMS
CoI	Social presence	Captures the extent to which learners experience a sense of connection, belonging, and interpersonal interaction	1. Active participation between instructors and learners in discussion forums or chats

		within online learning environments, despite physical separation.	2. Perceived sense of community and established safe space in the online class environment 3. Peer encouragement and support during both online and offline session
CoI	Cognitive presence	Describe the ways instructors sustain the learner's attention with teaching strategies that aligned the content and tasks within online learning environments	1. Effective use of digital tools to support understanding of the learning materials (e.g., recorded lecture, short videos, slides) 2. Facilitation of discussions that encourage exploration of ideas 3. Design activities that support application of concepts to real-world contexts
CRT	Cultural awareness	How instructors demonstrate awareness of students' diverse cultural backgrounds in online learning by acknowledging, anticipating, and responding to cultural differences in communication styles, learning expectations, and participation patterns.	1. Acknowledgement of students' diverse cultural backgrounds during instruction 2. Adapting different communication approaches to suit culturally diverse learners 3. Awareness to varying levels of participation or silence in online settings
CRT	Cultural relevance	Describes how instructors intentionally design and deliver online learning content, examples, and activities that are aligned with students' cultural contexts, making learning more meaningful and relatable.	1. Use of culturally relevant examples in teaching 2. Design of instructional materials that reflect cultural diversity 3. Encourage students to connect learning to their own cultural experiences into discussions and tasks
CRT	Cross cultural collaboration	Reflects how instructors facilitate and structure opportunities for students from diverse cultural backgrounds to interact, collaborate, and learn from one another within online environments to form a learning community	1. Design of group activities involving culturally diverse peers 2. Facilitation of inclusive interaction across cultural backgrounds in online discussions 3. Support for managing challenges in cross-cultural online collaboration

In total, the analysis process underwent several iterative cycles before arriving at the final themes. Initial coding was done primarily through in-vivo coding technique (Saldana, 2016) emphasizing on the actual words used by the participants to describe or express their experiences. General themes focused on clustering the initial coding into broader patterns, whereas the CoI-CRT sensitizing process reexamined how these themes intersect with the elements of CoI-CRT as guided by Table 4, leading to the finalized

themes for reporting the findings. An excerpt of codebook is presented below in Table 5 to illustrate how the data was thematized from coding to theme through the analytical steps applied in this study.

Table 5. Excerpts of codebook used in developing codes to theme

Initial codes (In-vivo coding)	General themes description	Value	Theme refinement (CoI sensitising process)	Naming finalized themes
Revisit lecture recordings	Self-paced and flexibility of online learning	+	Cognitive presence	Theme 1: Navigating the check and balance of online learning experiences
Pause and rewinds		+	Cognitive presence	
Can re-review difficult topic		+	Cognitive presence	
Can fall behind	Not staying focused	-	Cognitive presence	
Delaying revision		-	Cognitive presence	
Clear instructions	Conducive online learning environment	+	Teaching presence	
Friendly and welcoming		+	Teaching presence	
Hard to follow	Language barrier	-	Cognitive presence Social presence	
Mixed languages used		-	Social presence	
Difficult to coordinate group project	Lack of social engagement	-	Social presence	
Less engaging		-	Social presence	

4. Findings

The findings are organised into three overarching themes students' experiences of online learning, the factors that shaped their positive online learning experiences as well as the challenges they encountered while learning online. While these themes emerged inductively from the data, the Community of Inquiry (CoI) framework provided a useful interpretive lens to understand and discuss how aspects of teaching, social, and cognitive presence shaped their experiences when viewed together with CRT perspective.

4.1. Theme 1: Navigating the Check and Balance of Online Learning Experiences

The reported students' experiences while online learning reflected varying degrees of check and balance learning experiences. Flexibility and convenience were often emphasized as main merits of online learning, greatly forming most of the students' experiences. A lot of participants appreciated the capacity of revisiting recorded lectures and administrating their study schedules independently. One participant shared

"I liked being able to revisit lecture recordings to fully understand the material. It allowed me to pause, rewind, and really digest difficult concepts that I might have missed in a live class."

Additionally, flexibility in online learning also allowed some students to manage conflicting priorities, such as balancing ongoing internship with their studies. As another participant explained:

"Online classes meant I could be at my internship [place] and still not miss lectures, but it required me to be very organized. There were days when juggling both was incredibly stressful."

Participants consistently noted that online learning provided them with significant scheduling flexibility and control over pace. Many appreciated being able to listen to recorded lectures, revisit difficult content, or balance academic work with personal commitments. However, this flexibility also required them to impose self-discipline on themselves. Several students mentioned that without a fixed class structure or physical presence in the classroom, they needed to create their own routines to maintain continuity in their studies. This often involved scheduling designated study hours, planning weekly

tasks, and revisiting materials proactively in order to avoid falling behind. This aligns with research showing that online learning environments place greater responsibility on learners to self-regulate, manage time independently, and monitor their own understanding (Broadbent & Fuller-Tyszkiewicz, 2018).

Students' narratives indicated that while instructional design and teacher organisation supported these practices, much of the effort still fell on learners to monitor their progress and maintain regular engagement. The self-initiated routines they developed demonstrate a heightened awareness of the need for personal accountability when learning online. For instance, one participant admitted:

"Sometimes, knowing I can always watch the recording later made me delay reviewing the material. In the end, the backlog became overwhelming, and I struggled to catch up before exams."

The flexibility afforded by online learning was frequently highlighted by students as a positive aspect of their experience, yet this flexibility also intersected with deeper cognitive and instructional design dynamics. From a cognitive standpoint, flexibility allowed students to revisit lecture recordings, pause and replay difficult explanations, and engage with materials at their own pace—actions that supported their meaning-making processes and helped them manage the cognitive load of learning in a second language. Many students described using this ability to control the pace helped enhance their understanding, support their comprehension, and resolve uncertainties. This demonstrates how flexibility facilitated the iterative process of reflection and inquiry when learning online associated with cognitive presence.

At the same time, students' ability to benefit from this flexibility depended heavily on the clarity and organisation provided by instructors. When teaching presence was strong—through clear task instructions, well-structured learning materials, and accessible guidance—students felt better able to navigate the open-ended nature of online learning. Conversely, when instructions were unclear or course materials were dispersed across multiple platforms, the same flexibility became overwhelming, prompting students to rely more on personal strategies to stay on track. In a way, flexibility was experienced not only as a feature of online learning but as a condition shaped jointly by teaching and cognitive presence.

However, while flexibility was celebrated, many students expressed dissatisfaction with the limited interaction that is inherent in online learning. The lack of face-to-face communication often left students feeling disconnected. As expressed by one participant,

"Online interaction feels less engaging; I miss the spontaneous discussions we had in physical classes. The lack of eye contact and real-time reactions makes it hard to feel connected."

This loss of immediacy diminished the sense of belonging students typically derive from shared physical classroom spaces, as it did not extend the same in online shared spaces. It also affected their experiences in conducting group project, which was initially meant to foster collaboration and establish networking among their peers from local and/or international background. Instead, they reported more on how the experience was marred by logistical challenges, as informed in the following:

"In group projects, it was harder to coordinate because we couldn't just meet after class like before. Sometimes we had to rely on long message threads to make decisions, which was inefficient."

In some instance, language barriers further complicated interactions. Students described their struggles to follow discussions when lecturers or peers shifted between languages, causing them to second-guess whether they had understood instructions correctly or whether their interpretations of concepts aligned with the rest of the class. Studies have shown that such linguistic overload weakens students' ability to sustain cognitive presence in online learning because they must devote substantial mental effort to basic

understanding before they can meaningfully process content (Cao, Zhu, & Meng, 2021). This uncertainty often translated into hesitation: students reported holding back their questions, delaying responses, or remaining silent in group discussions out of fear of misunderstanding, mispronouncing terms, or giving an inaccurate answer. As one participant reported

“Sometimes I couldn’t follow discussions because of the mix of English and Malay terms. This made me hesitant to participate, fearing I might misunderstand or give the wrong response.”

Upon probing for example, the participant showed a Forum Discussion inside the university’s LMS, and narrated *“Like this, the first posting was in English, but discussion after is mixed with Malay. I can translate but not sure if I will disturb [interfere in the discussion]”*

Such experiences reveal that reduced social presence—combined with language uncertainty—intensified feelings of isolation and withdrawal in online learning spaces. Across the participants’ accounts, the lack of social presence stood out as the most significant gap in their online learning experiences. To address this, some participants indicated including more frequent instructor-driven discussions for opportunities to have more social presence in online environment. As one participant commented:

“If the professor could host informal Q&A sessions, it might make the environment feel more interactive.”

Overall, the students’ accounts illustrate how their online learning experiences were shaped by an interplay of flexible structure for higher autonomy, instructional clarity for better cognitive engagement, as well as more desirable social connection to provide better online learning experience. Viewed through the CoI framework, their experiences suggest that online learning environments were strongest in delivering content in flexible manner especially when presented in well-structured forms and interactive materials, but less effective in fostering interaction and deeper meaning-making, highlighting the need for a more balanced integration, especially the element of social presence, in online learning settings.

4.2. Theme 2: Enablers of Effective Online Learning Experiences

Under this theme, participants also described several conditions that helped them learn more effectively in the online environment. These “enablers” reflected features of the learning experience that were most appreciated by these students.

First and foremost, efficient instructor involvement was key to forming positive online learning experiences. This aligns with Garrison, Anderson, and Archer’s (2000) concept of teaching presence, where effective instructional design and clear organisation provide the foundation for purposeful, meaningful learning experience. For example, one participant noted

“Due to the professor’s rapid [quick] respond to our questions, I felt backed and less anxious about hard [tough] homework. It displayed that they cared about our advancement [progress in study].”

Then instructors responded promptly to questions or clarified confusing points during live sessions, students reported heightened motivation and reduced cognitive strain. Anderson et al. (2001) highlight that timely and continuous facilitation strengthens learners’ engagement by ensuring that instructional support is ongoing rather than episodic. For China international students, such responsiveness played an even more crucial role in reducing language-related uncertainties and ensuring they interpreted

instructions correctly. This responsiveness made the presence of instructors felt even in asynchronous spaces, reinforcing the perception that help was available when needed.

Students also shared that interactive teaching approaches played an important role in keeping them engaged and helping them understand course content more deeply. Activities such as breakout discussions, simulations, and case-based tasks encouraged them to move beyond passive listening and actively apply what they learned, as illustrated below:

“The simulated courtroom activities made the legal ideas come alive. They pushed us to think critically and use principles in actual scenes.”

“Our professor adopted case research and asked us to display solutions, which maintained us involved throughout the semester. It also assisted us in developing problem-solving abilities related to our real lives.”

These strategies supported cognitive presence while learning online, as they required students to make sense of ideas, analyse situations, and use concepts in meaningful ways. Recent studies show that online learning becomes more effective when instructors integrate authentic tasks and opportunities for problem-solving, which prompt learners to engage in deeper thinking and reflection (Fiock, 2020; Cleveland-Innes & Campbell, 2020). Students in this study similarly found that simulations and real-world cases made lessons more relatable and sustained their interest. This aligns with recent evidence suggesting that purposefully designed interactive tasks and instructional strategies help strengthen cognitive presence by promoting inquiry, application, and reasoning (Sadaf, Wu, & Martin, 2020).

What also emerged from the participants’ narratives was sensitivity to cultural differences as exhibited by some instructors during teaching. Teachers who knew cultural difference stood out as especially useful, illustrated as follow:

“When professors tried to know our cultural views, it felt like they truly cared about our study. For example, one professor modified instances to incorporate cases from Asia, which made the material more relatable.”

The above highlighted how teaching presence can intersect with the concept of culturally responsive teaching (CRT), which emphasise the importance of validating learners’ cultural identities and building instructional connections that resonate with their lived realities (Gay, 2010). Beyond providing clarity and organisation, students valued instructors who acknowledged and incorporated their cultural perspectives into course discussions and examples. Such efforts signalled acknowledgement for their backgrounds and presence in the class, and made learning materials feel more meaningful. When instructors adapted examples to include Asian contexts or invited students to share views shaped by their cultural experiences, students felt seen and understood. Within the CoI framework, these practices enriched teaching presence by strengthening both design and facilitation in ways that were culturally attuned rather than one-size-fits-all. They helped reduce the cultural distance that international students often feel and supported a more inclusive and personally relevant learning experience.

Lastly, as online learning environment is heavily mediated by the use of technology-based platform, strong digital platforms were also reported to be important in driving positive online learning experiences. Participants noted favourably on features like resource sharing, live discussions, and calendar functions for enhancing productivity. One participant observed:

“Microsoft Teams is comprehensive; it allows me to access resources, submit assignments, and engage in discussions seamlessly. The integration of all these functions saves a lot of time [when learning online].”

However, students also emphasized the need for ongoing training to maximize the potential of these platforms. When asked during LMS walkthrough, the participant illustrated some navigation that move back and forth between different courses and postings by weeks.

“There were so many functions I didn’t know how to use until halfway through the semester. More training sessions would be helpful, especially at the beginning of the [semester] term.”

These forms of support may act as a practical extension in considering aspect of teaching presence in online learning. When students received ample guidance on how to use online features provided, they reported feeling less anxious and more confident in engaging with course requirements. This echoes Martin, Sun & Westine’s (2020) argument that instructor support for technology, including clear instructions and troubleshooting guidance, directly enhances learners’ perceptions of effective online teaching. Ensuring learners can access and navigate course tools effectively when learning can extend and affirm the concept of teaching presence. Additionally, the strategic use of digital tools, when paired with clear instructions, can scaffold cognitive processes and help students engage more meaningfully with course content (Sadaf, Wu, & Martin, 2020).

4.3. Theme 3: Constraints on Effective Online Learning

As part of their shared experiences, the students also reported several constraints that made their online learning experience challenging and, at times, discouraging. While no one learning experience can ever be flawless, knowing about these constraints in online learning can serve as important lessons learned for future improvements in online teaching practices.

From their narratives, the data revealed the interplay of Chinese cultural perspectives within students’ online learning experiences. Many of the difficulties they described—such as hesitancy to speak, fear of misunderstanding instructions, and reluctance to interrupt or ask questions—were closely tied to Chinese cultural norms surrounding communication, learning, and classroom behaviour. For example, one participant reported

“It was hard to adapt to the open classroom style where students are expected to speak up. Back home, we were more used to lectures where the teacher led everything.”

Language and cultural constraints also posed significant challenges for many participants, particularly in adapting to local accents and unfamiliar teaching methods. As elaborated by the participants:

“Sometimes I couldn’t observe discussions due to the blend of English and Malay terms. It felt like I was always a procedure behind in insight.”

“In chat discussions, it’s difficult to tell the tone, and I sometimes miss explain what others represent [said]. This has resulted in awkward moments where I had to clarify my reactions.”

Chinese learners often come from educational environments that emphasize accuracy, respect for authority, and maintaining social harmony, which can discourage risk-taking or public expression of uncertainty (Heng, 2018). These values shaped how students responded to reduced social cues, language switching, and the unpredictable interaction patterns of online platforms. In the absence of clear guidance or immediate reassurance, students defaulted to culturally familiar strategies such as staying silent, observing before participating, or deferring to peers they perceived as more linguistically confident. As a result, the online environment amplified their tendencies toward being self-cautious,

indicating that their challenges were not only technical or pedagogical in nature but also deeply influenced by the socio-cultural influence they brought into the learning space.

Another constraint reported by the students is the technical issues faced when online learning. Technical difficulties, especially network unsteadiness, hampered their learning experiences as it can disrupted their cognitive presence during online learning. One participant mentioned:

“Sometimes the network lagged, and I missed core parts of the lecture. It was frustrating because it felt like I was missing out on crucial explanations.”

These disruptions affected more than just access, it also impacted their cognitive presence. It interfered with the students’ ability to process information and maintain continuity in their thinking in order to make sense of the course content. This aligns with studies showing that technical instability increases cognitive load and reduces learners’ capacity to engage in deeper inquiry or problem-solving during online lessons (Adedoyin & Soykan, 2023). Students also expressed the need for more reliable and accessible technical support or alternative resources to address the disruptions they frequently encountered during online sessions. These participants suggested:

“Having a devoted tech back team for students would make a great diversity [to the teaching practice]. They could provide real-time help during live meetings.”

“Professors should always upload lecture slides or offer alternative methods for access content in case of failed live session. It guarantees no one is left behind.”

Although technical difficulties may appear to fall outside the instructor’s control, their narratives hinted on how the element of teaching presence plays a significant role in mitigating the impact to their learning. When students encountered technical difficulties, what made the difference was not the technical failure itself but how instructors anticipated, responded to, and supported learners through these disruptions. Teaching presence involves instructional design, facilitation, and direction (Garrison, Anderson, & Archer, 2000), and in online contexts, this includes planning for technological breakdowns and providing clear alternatives to support multimodalities when learning online. In this sense, even though technological issues were unavoidable, strong teaching presence—through thoughtful organisation, contingency planning, and clear communication—helped buffer the disruptions and enabled students to continue learning effectively.

5. Discussions and Implications

Using the Community of Inquiry (CoI) framework as an interpretive lens, this study highlights how the three CoI elements —teaching presence, cognitive presence, and social presence—shaped Chinese international students’ experiences of online learning in Malaysian higher education. Although the themes emerged inductively, the CoI model provides a useful lens to sensitise their experiences from these three interrelated aspects.

Teaching presence emerged as a central determinant of students’ online learning experiences. The findings implied that teaching presence must be intentionally designed and actively maintained in online courses. When instructors provided prompt feedback and structured learning pathways, students reported feeling more confident and better supported. This aligns with research emphasising that strong instructional design, visible facilitation, and responsive communication form the backbone of effective teaching presence in online environments (Garrison, Anderson, & Archer, 2000; Fiock, 2020). Moreover, teaching presence should extend beyond content delivery to include proactive monitoring of students’ progress, timely clarification of misunderstandings, and the provision of alternative learning pathways when challenges arise. Such practices not only support students’ academic confidence but

also help stabilise their engagement and sustain learning momentum in online environments. From the perspective of CRT, intentionally designed instructions must include elements of cultural awareness, whereby instructors actively consider students' cultural backgrounds, language use, and prior learning experiences when structuring learning activities (Gay, 2010). In online settings, this is particularly important, as culturally informed behaviours—such as hesitation to speak or preference for indirect communication—may otherwise be misinterpreted as disengagement (Heng, 2018).

From the reported findings, cognitive presence was reported most positively when students engaged in interactive and authentic learning activities. Simulations, case-based learning, and breakout discussions prompted students to analyse problems, apply principles, and construct meaning—behaviours consistent with deeper inquiry processes described in recent CoI research (Cleveland-Innes & Campbell, 2020). A similar consideration can also be applied from a CRT perspective, as what Gay (2010) refers to as cultural awareness. When abstract concepts are being explained with familiar contexts, deeper meaning-making in the learning can be achieved.

However, cognitive presence can be highly susceptible to disruption when it comes to learning online. Technical challenges such as lagged lectures, an unstable internet may interrupt students' thought processes and made it difficult to follow complex explanations. These disruptions increased cognitive load and hindered sustained engagement, echoing findings that technical instability directly undermines learners' ability to maintain cognitive presence (Adedoyin & Soykan, 2023). Therefore, these findings implied that instructors should be mindful of the cognitive load imposed by online learning, particularly for multilingual learners. When students must simultaneously process content, manage technical issues, and navigate language demands, their capacity for deeper engagement may be reduced.

Among the three elements, social presence was the most vulnerable element reported in online learning. Students consistently reported feeling disconnected due to minimal interaction, absence of nonverbal cues, camera-off cultures, and limited opportunities for spontaneous, informal conversation with peers. These barriers align with research showing that online learning environments often struggle with disconnection and isolation, and lacking in verbal and non-verbal behaviors, thus limiting social presence (Richardson et al., 2017). For Chinese international students, cultural norms around communication—such as being cautious in public expression, fear of error, and deference to authority—coupled with multilingual language barrier, can heighten hesitation and withdrawal (Heng, 2018; Tu, 2010). As a result, it can have reduced their confidence leading to suppressed participation, and indirectly hamper their cognitive engagement when learning online. This suggests that instructors need to deliberately design opportunities for interaction, such as cross-cultural collaborations because in online learning environment, social presence cannot be assumed by default (Tu, 2010). This can be cultivating in support using the third principle of CRT, which calls for a learning culture that centered on diverse partnerships and group functioning where everyone helps one another and everyone contributes to mutual gain.

The implications of this study can also be considered in terms of immediate intervention in the instructional design, as well as long-term institutional strategies, particularly in strengthening the integration of CoI presences through culturally responsive approaches. Instructors may consider designing low-risk interaction opportunities, such as using chats and polls to support social presence among the students who hesitant to speak. Embedding culturally responsive teaching practices—such as structured peer discussions based on context relevance and low-risk participation activities that reward efforts and attempts—can help strengthen social presence and foster a greater sense of belonging in online learning environment. Clear and consistent communication can also help reduce uncertainty and enhance teaching presence. Suggestions may include providing weekly summaries of classes completed and list of things due or topics to be prepared for the following weeks.

Meanwhile, long-term strategic implications may include preparing professional development courses for instructors on exposure to the CoI-CRT principles and strategies. Institutional-level guidance on the CoI-CRT framework should be proposed to systematically embed teaching, social, and cognitive presence into course design rather than treating them as incidental elements.

6. Study Limitations and Recommendations for Future Research

This study was intentionally designed as an exploratory qualitative inquiry aimed to gain in-depth understanding of Chinese international students' online learning experiences from two intersecting perspectives from CoI and CRT. As such, the sample was limited to a small group of students from two public universities in Malaysia, allowing for rich, contextualised insights within the specified settings, but limiting the generalisation of the findings to other institutional or national contexts.

Future research can build on these findings by examining a wider range of institutions, including private universities or transnational programmes, and by purposively involving larger and more diverse samples of international students to reveal and analyse greater variation in the experiences reported. Mixed-methods or quantitative studies could also be employed to test the patterns identified here as variables, and to examine the relative strength of CoI three presences with CRT elements across different contexts. Such approaches would complement the present study by extending its exploratory insights and contributing to a statistically significance models connecting both the theoretical grounding of CoI and CRT together. In addition, a longitudinal study can also be considered to examine how these students' online learning experiences change over time, particularly in relation to the dynamic interplay of teaching, cognitive, and social presence. Such an approach would provide insights into how culturally responsive practices and students' engagement and participation develop across different stages of their learning journey.

7. Conclusion

Taken together, the findings of the study show that a balanced integration of teaching, cognitive, and social presence, when considered along with culturally responsive teaching (CRT), is critical for effective online learning, particularly for culturally diverse learners. The study highlights that China international students' online experiences are shaped not only by pedagogy and technology, but also by cultural perspectives, language barrier, and technical issues. Rather than viewing teaching, cognitive, and social presence as purely online instructional constructs, the findings illustrate how these dimensions are layered by cultural expectations, communication norms, and learners' prior experiences in their home countries. Positive experiences emerged when instructional clarity, cognitive engagement, and interpersonal connection were aligned; negative experiences often reflected weaknesses in instructional design in one or more of these CoI elements. In redesigning online learning using the CoI framework, this study imparts how the learners' cultural and contextual realities aligned well with the CRT principles. As online learning becomes a strategic feature of internationalised higher education, it is no longer sufficient to consider instructional design in isolation from cultural context. This study highlights how the integration of CoI and CRT offers a more holistic lens for understanding international students' experiences, emphasising that meaningful engagement emerges from the alignment of cultural responsiveness with teaching, cognitive and social presences.

References

- Adedoyin, O. B., & Soykan, E. (2023). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments*, 31(2), 863–875. <https://doi.org/10.1080/10494820.2020.1813180>
- Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001).
- Arbaugh, J. B. (2013). Does academic discipline moderate CoI-course outcomes relationships in online MBA courses? *The Internet and Higher Education*, 17, 16–28. <https://doi.org/10.1016/j.iheduc.2012.10.002>
- Bowen, G. A. (2006). Grounded Theory and sensitizing concepts. *International Journal of Qualitative Methods*, 5(3), 2-23. <https://doi.org/10.1177/160940690600500304>.
- Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- British Educational Research Association. (2018). *Ethical guidelines for educational research* (4th ed.). British Educational Research Association. <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018>
- Broadbent, J., & Fuller-Tyszkiewicz, M. (2018). Profiles in self-regulated learning and their correlates for online and blended learning students. *Educational Technology Research and Development*, 66, 1435 - 1455. <https://doi.org/10.1007/s11423-018-9595-9>
- Cao, C., Zhu, C., & Meng, Q. (2021). Chinese international students' coping strategies, social support resources in response to academic stressors: Does heritage culture or host context matter? *Current Psychology*, 40, 242–252. <https://doi.org/10.1007/s12144-018-9929-0>
- Castulo, N. J. (2025). Toward learning recovery: challenges of international students in China during COVID-19 online learning experience. *Quality Education for All*, 2(1), 100 – 113. <https://doi.org/10.1108/QEA-07-2024-0057>.
- Cleveland-Innes, M., & Campbell, P. (2020). *An introduction to distance education: Understanding teaching and learning in a new era* (2nd ed.). Routledge.
- Fiock, H. S. (2020). Designing a community of inquiry in online courses. *International Review of Research in Open and Distributed Learning*, 21(1), 135–153. <https://doi.org/10.19173/irrodl.v20i5.3985>
- Garrison, D. R. (2017). *E-learning in the 21st century: A framework for research and practice* (3rd ed.). Routledge.
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2–3), 87–105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Garrison, D. R., & Arbaugh, J. B. (2007). Researching the Community of Inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3), 157–172. <https://doi.org/10.1016/j.iheduc.2007.04.001>
- Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). Teachers College Press.

- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>.
- Heng, T. T. (2018). Different is not deficient: Contradicting stereotypes of Chinese international students in U.S. higher education. *Studies in Higher Education*, 43(1), 22–36. <https://doi.org/10.1080/03075079.2016.1152466>
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Sage Publications.
- Jones, E. (2013). Internationalisation and employability: The role of intercultural experiences in the development of transferable skills. *Public Money & Management*, 33(2), 95–104. <https://doi.org/10.1080/09540962.2013.763416>
- Kathirveloo, P. (2024). Assessing institutional readiness for blended learning management and implementation in Malaysian higher education. *Management Research Journal*, 13(2), 55–67. <https://doi.org/10.37134/mrj.vol13.2.5.2024>
- Khoo, E., & Huo, X. (2022). The efficacy of culturally responsive pedagogy for low-proficiency international students in online teaching and learning: A Canadian experience with an online reading–writing program. *Journal of Teaching and Learning*, 16(2), 67–85. <https://doi.org/10.22329/jtl.v16i2.7022>
- Knight, J. (2004). Internationalisation remodeled: Definition, approaches, and rationales. *Journal of Studies in International Education*, 8(1), 5–31. <https://doi.org/10.1177/1028315303260832>
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465–491.
- Leask, B. (2015). *Internationalizing the Curriculum*. Routledge.
- Liu, D., & Morgan, W. J. (2017). Chinese students overseas: Choice of destination. In W. J. Morgan, Q. Gu, & F. Li (Eds.). *Handbook of education in China* (pp. 442 – 467). Elgar Publishing. <https://doi.org/10.4337/9781783470662.00031>
- Mahani, S. (2023). Applying culturally relevant pedagogy to online learning. *Cultural and Pedagogical Inquiry*, 15(2), 23–34. <https://doi.org/10.18733/cpi29708>
- Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, 159, 104009. <https://doi.org/10.1016/j.compedu.2020.104009>
- Ministry of Higher Education Malaysia. (2025). *Statistic of Higher Education*. <https://mohe.gov.my/muat-turun/statistik/2024-4/1702-bab-1-makro-2024-update-pdf?highlight=WyJzdGF0aXN0aWsiXQ==>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *Internet and Higher Education*, 14(2), 129–135. <https://doi.org/10.1016/j.iheduc.2010.10.001>

Muniisvaran, K., Jose, F.T., Anusia, P. K. (2025). Technology-enhanced learning in higher education institutions in Malaysia. *International Journal of Advanced and Applied Sciences*, 12(1), 220 – 225. <https://doi.org/10.21833/ijaas.2025.01.021>

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health, Education, and Welfare <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>

Ng, D. T. K., & Fang, X. (2023). Inequity issues in online learning of Chinese cross-border students under the COVID-19 pandemic. *Educational Technology & Society*, 26(4), 123–135. <https://www.jstor.org/stable/48747525>

Organisation for Economic Co-operation and Development (OECD) (2025), What are the key trends in international student mobility? *Education Indicators in Focus*, 88. OECD Publishing, <https://doi.org/10.1787/2a423a76-en>.

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publication.

Rahim, N. A. A. (2021). Adjustment experiences of international graduate students to life in the United States. *Pertanika Journal of Social Sciences & Humanities*, 29(4), 2243–2262. <https://doi.org/10.47836/pjssh.29.4.08>

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, 71, 402–417. <https://doi.org/10.1016/j.chb.2017.02.001>

Sadaf, A., Wu., T., & Martin, F. (2021). Cognitive presence in online learning: A systematic review of empirical research from 2000 to 2019. *Computers and Education Open*, 2 (December 2021), 100050. <https://doi.org/10.1016/j.caeo.2021.100050>

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage Publications.

Triandis, H. C. (1995). *Individualism and collectivism*. Westview Press.

Tu, C. H. (2010). How Chinese perceive social presence: An examination of interaction in online learning environment. *Educational Media International*, 38(1), 45–60. <https://doi.org/10.1080/09523980010021235>

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>

Wei, Y. (2024). Application of blended learning in the Intercultural Communication Course in English in Chinese universities. In Chowdhury, R., Tuan, H.A. (Eds.) *Engaging with Australasia: Comparative Research on ELT and English Teacher Education*. Palgrave Macmillan. https://doi.org/10.1007/978-981-97-4038-3_7

Zhou, Y., & Todman, J. (2009). Patterns of adaptation of Chinese postgraduate students in the United Kingdom. *Journal of Studies in International Education*, 13(4), 467–486.
<https://doi.org/10.1177/1028315308317937>

Leadership Development in United States Higher Education: A Systematic Review of Integrated Approaches for University Students

ABSTRACT

Authors: Zurina Ahmad Saidi, and Mohd Faiq Abd Aziz

Affiliation: Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:

mohdfaiq@upm.edu.my

Received: 31/12/2025 |

First Revision: 14/1/2026 |

Second Revision: 15/4/2026 |

Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

Leadership skill development has become a strategic priority in higher education as universities seek to prepare graduates for increasingly complex professional and societal roles. Although research on leadership development has expanded, existing studies often examine specific approaches in isolation, limiting an integrated understanding of how leadership capacities are cultivated among university students. This study systematically synthesises leadership development strategies within United States higher education to identify coherent patterns that support student leadership learning and career readiness. Guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, a systematic search of the Scopus database was conducted in December 2025. This review focuses exclusively on studies conducted within United States higher education contexts. Following screening and eligibility assessment, 59 peer-reviewed journal articles were included in the final synthesis. Thematic analysis identified six interconnected dimensions of leadership development: experiential learning, student-led co-curricular engagement, formal leadership education, mentored leadership development, professional and soft skill development, and research and innovation-based learning. The findings suggest that leadership development is strengthened when approached as an integrated, student-centred process that aligns curricular, co-curricular, relational, and inquiry-based learning environments. An integrated leadership development framework is proposed to inform comprehensive leadership initiatives in higher education.

Keywords: Leadership development, higher education, university students, experiential learning, co-curricular leadership, systematic review, mentoring, leadership education

1. Introduction

Leadership development has become a central priority in higher education as universities are increasingly expected to prepare graduates who are not only academically competent but also capable of leading within complex professional and societal environments. Rapid changes in the global labour market, coupled with rising expectations for graduate employability, have intensified the demand for leadership skills such as communication, adaptability, problem solving, and collaboration. As a result, leadership development is no longer viewed as an optional or elite activity but as an essential outcome of the university experience (Knight & Yorke, 2004; Dugan, 2017).

Existing research demonstrates that leadership learning within higher education is multifaceted and extends beyond classroom instruction. Studies emphasise experiential learning, student led co-curricular involvement, formal leadership education, mentoring relationships, professional and soft skill development, as well as research and innovation based learning as key pathways through which leadership capacity is cultivated (Astin & Astin, 2000; Komives et al., 2005; Brungardt et al., 2006; Crisp & Cruz, 2009; Healey & Jenkins, 2009). Collectively, these studies suggest that leadership development is a developmental and contextual process shaped by both structured and informal learning experiences.

However, despite the growing body of scholarship on leadership development, the literature remains conceptually fragmented. Many studies examine individual strategies in isolation, focusing separately on classroom-based leadership education, co-curricular participation, mentoring relationships, or experiential engagement (Dugan & Komives, 2010; Seemiller, 2016; Guthrie et al., 2017). While these contributions provide valuable insights, limited attention has been given to how these diverse approaches intersect, complement one another, or collectively contribute to coherent leadership learning pathways within higher education institutions. The absence of integrative synthesis constrains the development of comprehensive frameworks capable of guiding holistic leadership initiatives across curricular and co-curricular domains.

In response to this gap, the present study systematically examines leadership skill development strategies among university students by synthesising evidence across six interconnected themes. By drawing on empirical studies within higher education, this study seeks to provide a comprehensive understanding of how leadership skills are developed and how these skills contribute to student's future career readiness. Through this integrative approach, the study aims to inform educators, policymakers, and higher education institutions seeking to design effective and inclusive leadership development initiatives.

Leadership development in higher education is increasingly recognised as a key contributor to graduate employability and long term career success. Research demonstrates that leadership skills developed during university studies, such as teamwork, communication, critical thinking, and self-awareness, are closely aligned with the competencies demanded by employers and professional environments (Robles, 2012; Jackson, 2014; Heckman & Kautz, 2012). Leadership learning is therefore not confined to preparing students for formal leadership positions but plays a broader role in enhancing students capacity to navigate complex workplace and societal challenges. Existing studies indicate that effective leadership development is supported through a combination of experiential learning, student led co-curricular engagement, structured leadership education, mentoring support, professional skill development, and research based learning environments (Komives et al., 2006; Owen et al., 2009; Allen & Hartman, 2008; Packard, 2016; Fung, 2017).

Given these considerations, there is a need for systematic synthesis that consolidates existing evidence and clarifies how various leadership development strategies collectively support student learning and professional readiness. Addressing this gap, the present study is guided by the following research question:

What leadership development strategies are identified in higher education literature, and how do these approaches collectively contribute to university students' leadership skill development and career readiness?

Accordingly, this study systematically synthesises leadership development strategies in higher education and examines how these interconnected approaches support the cultivation of leadership capacity among university students. By integrating diverse leadership development pathways within a unified analytical framework, the study contributes to a more coherent conceptualisation of student-centred leadership learning and proposes an integrated model to inform comprehensive institutional leadership initiatives.

2. Methodology

2.1. Research Design

This study adopted a systematic literature review design guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to ensure transparency, procedural clarity, and reproducibility. A systematic approach was employed to identify, screen, evaluate, and synthesise empirical evidence concerning effective strategies for leadership skill development among university students in higher education. The use of PRISMA enabled structured documentation of the identification, screening, eligibility, and inclusion stages, thereby enhancing procedural clarity and minimising selection bias.

2.2. Database Selection and Search Strategy

The literature search was conducted in December 2025 using the Scopus database. Scopus was selected due to its comprehensive coverage of peer-reviewed journals across social sciences and education-related disciplines, as well as its advanced search functionalities that allow precise Boolean structuring and field-specific filtering. To ensure conceptual alignment with the research objective, the search was restricted to the TITLE-ABS-KEY fields. The following Boolean search string was applied:

Table 1. Search Strings

Searching keyword
(TITLE-ABS-KEY("leadership") AND TITLE-ABS-KEY("skill" OR "competence" OR "capability") AND TITLE-ABS-KEY("development" OR "training") AND TITLE-ABS-KEY("higher education" OR "university students" OR "college students")) AND TITLE-ABS-KEY("strategies" OR "approaches" OR "methods"))

The initial search yielded 1477 records. To refine the search results and enhance disciplinary relevance, filters were applied to include only peer-reviewed journal articles (document type: article) written in English. Following the application of these filters, 275 records remained for screening. While Scopus was selected due to its comprehensive coverage of peer-reviewed literature in social sciences and education, it is acknowledged that relevant studies indexed in ERIC, PsycINFO, and Web of Science may not have been captured. The use of a single database was intended to ensure consistency and manageability of the review process, although it may limit the breadth of coverage and introduce potential publication bias.

2.3. Inclusion and Exclusion Criteria

The review was limited to studies conducted in the United States to ensure contextual consistency across higher education systems, particularly in relation to institutional governance structures, the maturity of co-curricular leadership ecosystems, and the integration of student affairs and academic development frameworks. The United States higher education context is characterised by well-established leadership

development infrastructures, including structured co-curricular programmes, mentoring systems, and experiential learning initiatives, which provide a relatively comparable basis for synthesis. However, this restriction may limit the transferability of findings to other national contexts, as leadership development practices are shaped by cultural, institutional, and policy differences. As such, the findings of this review should be interpreted as being strongly informed by United States higher education contexts, while offering indicative insights that may require contextual adaptation in other settings.

2.4. Article Selection Process (PRISMA)

The study selection process followed the four-stage PRISMA procedure: identification, screening, eligibility, and inclusion. During the identification stage, 1477 records were retrieved from Scopus. After applying document type and language filters, 275 records remained. These 275 articles underwent title and abstract screening to assess relevance to leadership skill development strategies within higher education settings. At this stage, 163 articles were excluded for failing to meet the inclusion criteria. The remaining 112 articles progressed to full-text assessment for eligibility. Each article was evaluated for methodological suitability, empirical focus, and alignment with leadership development strategies targeting university students. During this stage, 53 articles were excluded due to conceptual misalignment, insufficient empirical evidence, or lack of explicit focus on leadership development strategies. Following full-text assessment, 59 studies met all inclusion criteria and were included in the final synthesis. The selection process is illustrated in the PRISMA flow diagram. Title and abstract screening was conducted by the authors, with a subset of studies independently double-screened to ensure consistency. Full-text screening was subsequently carried out, and any discrepancies were discussed and resolved through consensus among the authors.

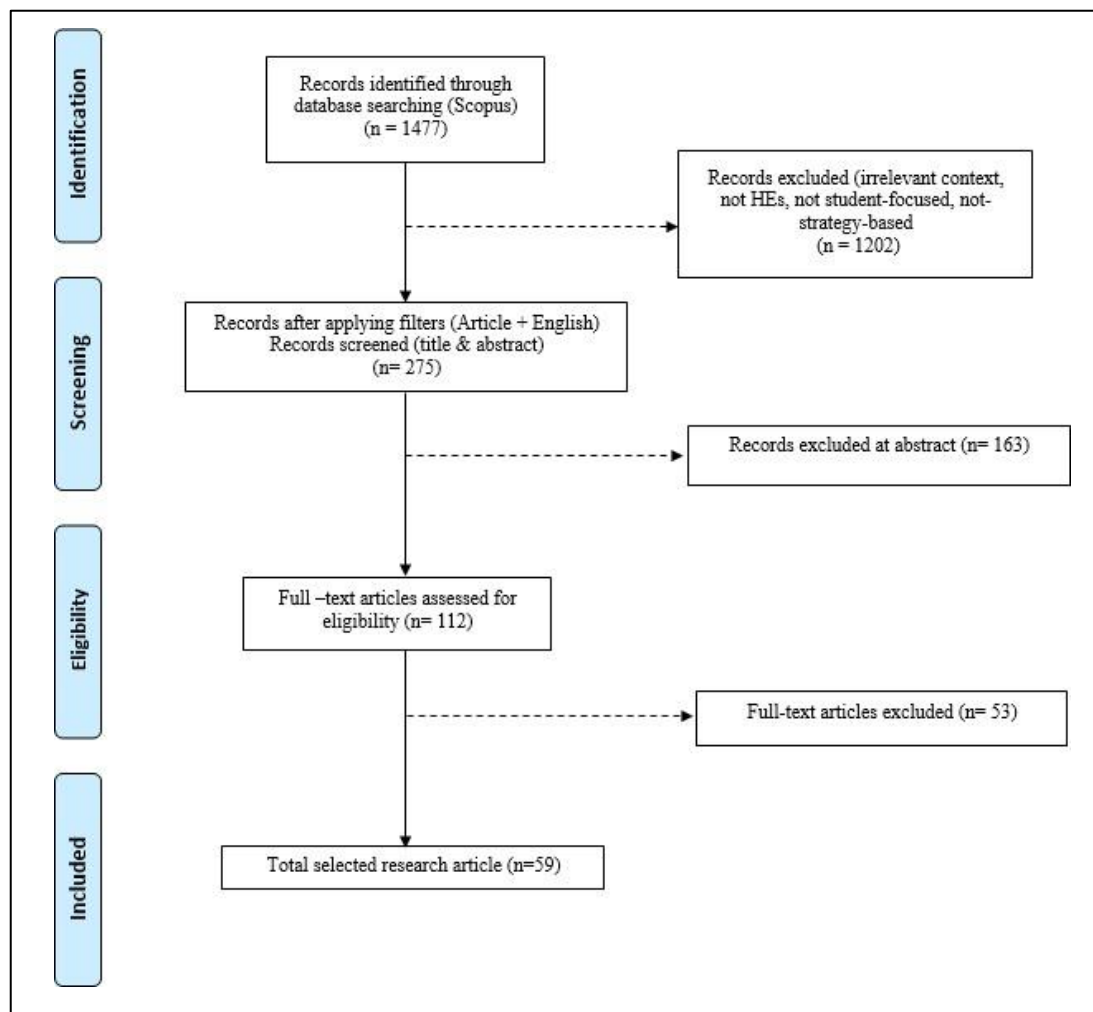


Figure 1. PRISMA Flow Diagram of the Article Selection Process

To enhance transparency and reproducibility of the screening process, the reasons for exclusion at each stage are summarised in Tables 2.

Table 2: Reason for Exclusion During Initial Screening, Abstract Screening Stage and Full Text Exclusion

Stage	Reason	Number	Total
Initial screening	• Not related to leadership development • Not focused on higher education context • Not student-focused (e.g., organisational leadership) • Not aligned with leadership development strategies • Duplicate or irrelevant records		1202
Abstract screening	• Insufficient focus on leadership development outcomes • Not empirical study • Weak relevance to research objectives • Conceptual or theoretical only (no data) • Incomplete or unclear findings		163
Full-text screening	• Not aligned with leadership development focus • Not conducted in higher education context • Not conducted in United States • Insufficient methodological rigor • Insufficient data for analysis	18 12 10 7 6	53

2.5. Data Extraction and Synthesis

A structured data extraction process was implemented to ensure systematic documentation and comparability across studies. A standardised data extraction form was developed using Microsoft Excel to capture key study characteristics, including author(s), year of publication, research design, sample characteristics, institutional context, leadership development strategy examined, and principal findings related to leadership skill development. The data extraction process was conducted by the authors and cross-checked to ensure consistency and accuracy. Any discrepancies identified during extraction were discussed and resolved through consensus.

Thematic synthesis was employed to analyse the extracted data. An inductive coding process was undertaken, in which study findings were examined line-by-line to generate initial codes. This process resulted in the identification of a broad set of preliminary codes capturing recurring concepts across the included studies. Through iterative comparison and refinement, these initial codes were progressively grouped into higher-order categories. Following two rounds of coding and thematic consolidation, the codes were synthesised into six overarching themes representing key leadership development pathways within higher education: experiential leadership learning; student-led co-curricular leadership; formal leadership education; mentored leadership development; professional and soft skill development; and research and innovation-based leadership learning.

Microsoft Excel was used to support the organisation and categorisation of codes during the analysis process. Where necessary, analytical discussions were conducted between the authors to ensure conceptual clarity and consistency in theme development. This iterative and collaborative approach enhanced the transparency and reliability of the thematic synthesis.

During the data organisation and synthesis process, AI-assisted tools (ChatGPT and Elicit) were used to support the structuring of extracted information and the identification of potential thematic patterns. Elicit was used to assist in identifying potential patterns and organising extracted data into preliminary thematic groupings. These outputs were not used directly but were critically reviewed, compared against the original studies, and substantially refined by the authors. All final coding decisions and thematic interpretations were determined solely by the authors. These tools were utilised as supportive analytical

aids rather than as primary sources of interpretation. All AI-generated outputs were critically reviewed, validated against the original source materials, and refined by the authors to ensure accuracy, consistency, and conceptual integrity.

2.6. Consideration of Methodological Quality

To enhance methodological transparency and address potential risk-of-bias concerns, a structured qualitative assessment of methodological quality was incorporated into the review process. Although a formal numerical scoring system was not applied, each included study was evaluated based on key criteria, including clarity of research design, appropriateness of methodological approach, adequacy of sample size, transparency of data collection procedures, and alignment between findings and conclusions. Based on these criteria, studies were categorised into three levels of evidence strength: strong, moderate, and limited. Studies classified as strong typically employed robust research designs (e.g., quasi-experimental, longitudinal, or large-scale quantitative studies) with clear methodological reporting. Moderate studies included well-conducted qualitative or mixed-methods research with adequate transparency, while limited studies were generally small-scale, single-institution, or descriptive program evaluations with less methodological detail.

This classification informed the interpretive weighting of evidence during thematic synthesis, allowing greater emphasis to be placed on studies with stronger methodological rigour. The evidence strength classification is reflected in the synthesis tables to provide readers with a clearer understanding of the robustness of the underlying evidence base. This approach ensures that the review maintains inclusivity across diverse study designs while enhancing transparency in how evidence is interpreted and synthesised.

3. Results and Discussion

3.1. Theme 1: Experiential Leadership Learning

The literature consistently positions experiential leadership learning as a fundamental pathway through which leadership capacity is developed among university students. Rather than relying on abstract or purely classroom-based instruction, leadership learning is understood to emerge through direct engagement in real and meaningful contexts. Early conceptual contributions emphasise leadership as a socially embedded and developmental process, shaped by students' active participation in institutional and community environments. Through involvement, interaction, and responsibility, students gradually construct their understanding of leadership, highlighting that leadership learning is rooted in experience and social practice rather than passive knowledge acquisition (Astin & Astin, 2000; Komives et al., 2005).

Empirical evidence further reinforces the importance of practising leadership in authentic settings. Studies show that students who take on leadership roles within university contexts develop higher levels of leadership competence, confidence, and efficacy as they are required to make decisions, coordinate with others, and respond to real challenges (Dugan & Komives, 2010). Experiential leadership programmes that incorporate group problem solving, applied projects, and real-world leadership tasks are found to promote deeper and more sustained learning outcomes compared to purely instructional approaches (Eich, 2008). These findings suggest that experiential learning is particularly effective because it allows leadership behaviours to be enacted, tested, and refined through practice.

Reflection emerges as a critical element that transforms experience into meaningful leadership learning. Experiential pedagogy is most effective when students are supported in reflecting on their actions, challenges, and outcomes, enabling them to interpret experiences and adapt their leadership approaches accordingly (Roberts, 2012). Leadership development is therefore viewed as a dynamic process shaped by challenge, feedback, and reflection, rather than a fixed set of skills acquired at a single point in time (Day et al., 2014). Reflective engagement also supports the development of leadership identity, as

students internalise leadership values and gain greater self-awareness through experiential learning (Owen, 2016).

Experiential leadership learning also plays a vital role in bridging theory and practice across diverse learning environments. High-impact practices such as service learning, internships, applied leadership projects, and outdoor or adventure-based programmes connect leadership concepts to real-world application, strengthening students' confidence and practical understanding of leadership (Guthrie & Jenkins, 2018; Priest & Gass, 2018). Engagement in community-based and experiential contexts further exposes students to complex social and organisational challenges, fostering adaptability and problem-solving skills (Soria et al., 2015). At the institutional level, leadership development is most effective when universities intentionally embed experiential engagement across the student experience, rather than treating it as an isolated or optional component (Kezar & Holcombe, 2019). Critical service-learning approaches further extend experiential leadership learning by linking leadership practice with social responsibility and community problem solving (Mitchell et al., 2015).

Overall, there is strong consensus across conceptual and empirical studies that experiential leadership learning represents a core pathway for leadership development in higher education. Leadership emerges through direct experience, is strengthened through reflection, and is deepened when theory is meaningfully connected to practice. By engaging students in authentic, challenging, and reflective learning environments, experiential approaches support leadership identity development, skill acquisition, and adaptive capacity, underscoring their central role in preparing students for leadership in complex real-world contexts. Importantly, experiential leadership learning does not operate in isolation but often intersects with co-curricular, mentoring, and research-based environments. This suggests that experiential engagement functions as a foundational mechanism through which other leadership development strategies gain practical meaning and developmental depth.

The evidence within this theme is predominantly derived from qualitative and program-based studies, with limited experimental validation, suggesting that while experiential learning is widely recognised as foundational to leadership development, the strength of causal evidence remains relatively constrained.

3.2. Theme 2: Student-Led Co-Curricular Leadership

Student-led co-curricular leadership is consistently identified as a significant context for leadership development, particularly through forms of engagement that extend beyond formal academic settings. Leadership learning in these spaces is closely tied to students' active participation in campus life, where leadership emerges through interaction, shared responsibility, and collective decision making. Student involvement in organisations and activities provides opportunities for leadership to be experienced as a lived practice rather than a theoretical concept, reinforcing the idea that leadership development is most effective when students are directly involved in shaping their own learning environments (Astin, 1999; Kezar & Moriarty, 2000).

Empirical research demonstrates that co-curricular settings offer important opportunities for students to practise leadership by organising activities, coordinating initiatives, and leading peers. Leadership roles within student organisations and governance structures allow students to engage with real responsibilities, strengthening leadership competence through experience rather than instruction alone (Zimmerman-Oster & Burkhardt, 1999; Dugan & Komives, 2007). Co-curricular leadership learning is driven by collaboration, accountability, and peer interaction, which together create conditions for the development of practical leadership skills and confidence (Haber, 2011).

A defining feature of co-curricular leadership learning is its informal and authentic nature. Student organisations often function as communities of practice in which leadership is shared, negotiated, and enacted through everyday interactions rather than formal hierarchies (Komives et al., 2006). These environments allow students to experiment with leadership roles in meaningful yet relatively low-risk contexts, supporting leadership identity development over time. Through trial, reflection, and social

engagement, students are able to explore different leadership approaches and develop a deeper understanding of themselves as leaders (Owen et al., 2009).

Beyond individual skill development, student-led co-curricular engagement also contributes to broader civic and social outcomes. Participation in student-led initiatives has been shown to enhance leadership confidence, interpersonal competence, and a sense of social responsibility (Saidi et al., 2025; Soria et al., 2012). Leadership and representation roles further support students' civic awareness and engagement with institutional decision making, positioning leadership as both a personal and collective endeavour (Luescher-Mamashela, 2013). Sustained involvement in co-curricular leadership activities is therefore closely associated with positive leadership outcomes, particularly when students are actively involved in leading peers and shaping shared goals (Dugan, 2017).

Overall, the literature highlights student-led co-curricular leadership as a vital pathway for leadership development in higher education. Through active participation in student-initiated activities, students learn to organise, collaborate, and lead in authentic and informal settings that complement formal learning. These experiences foster leadership confidence, identity development, and transferable skills, underscoring the central role of co-curricular engagement in cultivating leadership capacity among university students. While co-curricular engagement offers authentic leadership practice, its developmental impact appears strongest when intentionally connected to reflective or instructional components. This highlights the importance of institutional coordination in aligning co-curricular leadership opportunities with broader leadership learning objectives.

The evidence supporting this theme is largely based on cross-sectional and self-reported studies, with limited longitudinal or experimental designs, indicating that observed positive outcomes may be influenced by self-selection effects rather than direct causal impact.

3.3. Theme 3: Formal Leadership Education

Formal leadership education is commonly described in the literature as a deliberate and structured approach to cultivating leadership capacity within higher education. Rather than emerging incidentally, leadership learning in this context is intentionally designed through curriculum planning and instructional strategies. Leadership is treated as an area of study that can be systematically taught, learned, and assessed, distinguishing formal leadership education from more informal or experiential approaches (Brungardt et al., 2006). Credit-bearing leadership courses play an important role in this process by providing coherent learning pathways that introduce students to leadership concepts, models, and competencies in a structured manner (Jenkins, 2012).

Classroom-based leadership courses contribute significantly to students' understanding of leadership by offering conceptual clarity and theoretical grounding. Through guided instruction and structured learning activities, students are able to engage critically with leadership theories and examine how these frameworks apply to practice (Rosch & Caza, 2012). Structured classroom environments also allow for the intentional integration of discussion, reflection, and applied exercises, reinforcing students' theoretical understanding and supporting deeper intellectual engagement with leadership concepts (Allen & Hartman, 2008). These approaches suggest that formal leadership education provides a level of conceptual organisation that may not always emerge through experience alone.

Curriculum-based leadership programmes further strengthen leadership development by aligning learning objectives, pedagogical strategies, and assessment practices. Leadership education frameworks embedded within academic curricula promote consistency and coherence in learning outcomes, ensuring that leadership development is intentional rather than incidental (Komives et al., 2013). Competency-based approaches extend this alignment by clearly defining leadership skills and developmental goals, allowing curricula to be designed around measurable and developmentally appropriate outcomes (Seemiller, 2016). Well-structured leadership programmes have been shown to enhance students' leadership learning by providing clarity, progression, and purposeful assessment (Guthrie et al., 2017).

Formal leadership education is also understood as a developmental process that unfolds over time through carefully sequenced instruction. Leadership learning is supported by curriculum design that recognises development as longitudinal, enabling students to build leadership knowledge progressively across courses and academic experiences (Day et al., 2014). When leadership training modules are systematically integrated into academic programmes, they contribute more effectively to leadership growth and application (Priest & Thorpe, 2017).

Academic leadership courses further strengthen leadership understanding by offering structured opportunities for both knowledge acquisition and practical application within classroom contexts (Boatman et al., 2019). However, formal leadership education alone may be insufficient without experiential reinforcement, suggesting that structured instruction and applied practice must function in complementary rather than isolated roles. Overall, the literature reflects strong agreement that formal leadership education plays a critical role in leadership development through structured and intentional instruction. By providing systematic knowledge, conceptual frameworks, and clearly defined learning pathways, formal courses and curriculum-based programmes strengthen students' theoretical understanding of leadership. When leadership education is embedded within academic curricula, higher education institutions create purposeful learning environments that support consistent, reflective, and conceptually grounded leadership development.

Compared to other themes, this area includes a relatively stronger methodological base, with some quasi-experimental and curriculum-based evaluations; however, the overall evidence remains uneven and context-dependent across institutions.

3.4. Theme 4: Mentored Leadership Development

Mentored leadership development is widely acknowledged as a significant pathway through which leadership capacity is cultivated in higher education. Rather than focusing solely on skill transmission, mentoring emphasises relational learning supported by guidance, dialogue, and ongoing support. Mentoring is commonly understood as a multidimensional developmental process that provides students with academic, psychosocial, and career-related support, all of which contribute to leadership growth (Crisp & Cruz, 2009). Similarly, mentoring has been described as an intentional relationship that facilitates learning through interpersonal exchange and guided reflection, positioning it as a structured yet relational approach that complements both formal leadership education and experiential learning (Allen & Eby, 2007).

A key strength of mentoring lies in the feedback, role modelling, and encouragement provided by mentors. Through close interaction, mentors model leadership behaviours and values that students observe and gradually internalise, contributing to the development of leadership identity (Komives et al., 2006). Mentoring relationships also enhance leadership outcomes by reinforcing effective leadership behaviours and offering personalised feedback tailored to individual experiences (Dugan & Komives, 2010). Coaching-oriented mentoring further supports learning by creating opportunities for students to reflect on leadership challenges with more experienced individuals, deepening understanding through guided dialogue (Dominguez & Hager, 2013).

Mentoring is also strongly associated with the development of leadership confidence and a more coherent sense of leadership identity. By providing affirmation, encouragement, and constructive feedback, mentoring supports students' belief in their leadership abilities and sustains motivation over time (Murphy & Johnson, 2011). This supportive function is particularly important during periods of transition and uncertainty within higher education, where mentoring can help emerging leaders navigate challenges and make sense of their leadership experiences (Packard, 2016). Intentional mentoring relationships therefore play a crucial role in fostering confidence and self-efficacy in leadership development (Johnson, 2016).

Beyond individual outcomes, mentoring supports leadership development as a continuous and evolving process across students' educational journeys. Adult guidance and sustained mentor engagement have been shown to influence leadership readiness and capacity, particularly when mentoring is embedded within broader developmental structures (Campbell et al., 2012). Reviews of mentoring models in higher education further highlight the value of long-term mentoring relationships in providing continuity, developmental alignment, and ongoing support that extend beyond isolated learning moments (Gershenfeld, 2014). These findings suggest that mentoring functions not as a one-time intervention, but as an ongoing mechanism that supports leadership growth over time.

Overall, there is strong agreement in the literature that mentored leadership development plays a vital role in cultivating leadership capacity among university students. Leadership develops through guided relationships that offer feedback, role modelling, and encouragement, while simultaneously strengthening confidence and leadership identity. By supporting continuous development across different stages of students' educational experiences, mentoring complements experiential, co-curricular, and formal leadership learning, affirming its importance as a relational and sustained pathway for leadership development in higher education. Mentoring therefore functions as a relational bridge that connects experiential, co-curricular, and formal learning contexts, reinforcing leadership identity development across multiple domains.

The evidence within this theme is primarily qualitative and relational in nature, with limited large-scale empirical validation, suggesting that while mentoring is conceptually significant, its measurable impact on leadership outcomes requires further rigorous investigation.

3.5. Theme 5: Professional and Soft Skill Development

The literature consistently identifies professional and soft skill development as a central outcome of leadership development within higher education. Leadership is commonly conceptualised as a competency-based construct that extends beyond positional roles to include a broad range of professional capabilities, such as communication, collaboration, and self-management (Seemiller, 2013). Soft skills are therefore viewed not as supplementary attributes, but as essential leadership outcomes that enable individuals to function effectively in organisational and professional contexts (Robles, 2012). From this perspective, leadership development serves as an important mechanism through which higher education cultivates skills that go beyond disciplinary or technical knowledge.

A strong emphasis is placed on communication, teamwork, and interpersonal competence as core components of leadership learning. Employability and professional capability are closely linked to students' ability to communicate clearly, work collaboratively, and engage effectively with others (Saidi & Abd Aziz, 2025; Knight & Yorke, 2004). Employers consistently value these attributes, particularly skills related to teamwork, adaptability, and interpersonal interaction, which are often developed through leadership-related learning experiences rather than traditional academic instruction alone (Andrews & Higson, 2008). Leadership learning environments provide practical opportunities for students to practise these skills through interaction, collaboration, and shared responsibility, strengthening their capacity for effective communication and relationship building (Crosbie, 2005).

Professional and soft skills are also widely recognised as enhancing both leadership effectiveness and employability. Evidence suggests that non-cognitive and soft skills play a critical role in long-term professional success, often exerting a stronger influence than purely technical abilities (Heckman & Kautz, 2012). Leadership engagement within higher education contributes significantly to employability by fostering transferable skills that align closely with workplace expectations, including problem solving, teamwork, and interpersonal effectiveness (Jackson, 2014). Students who actively participate in leadership activities consistently demonstrate stronger development of these professional skills, highlighting the practical value of leadership learning experiences (Dugan, 2017). This positioning reframes leadership development not merely as preparation for formal leadership roles, but as the cultivation of adaptive professional capacity across diverse career trajectories.

Emotional and social competencies further reinforce the connection between leadership learning and professional skill development. Emotional intelligence has been strongly associated with effective leadership, particularly through qualities such as self-awareness, empathy, and relationship management (Goleman, 2000). Leadership training initiatives have been shown to support the development of these competencies, equipping students with soft skills that are transferable across a wide range of professional contexts (Laker & Powell, 2011). Collectively, these findings suggest that leadership learning provides a critical foundation for preparing students to navigate complex interpersonal and organisational demands.

Overall, the literature clearly demonstrates that leadership development plays a vital role in supporting professional and soft skill growth among university students. Through leadership learning, students develop communication, teamwork, and interpersonal skills that enhance leadership effectiveness and employability. These transferable competencies equip graduates with the professional capabilities needed to succeed across diverse career pathways, underscoring the importance of professional and soft skill development as a fundamental dimension of leadership learning in higher education.

The supporting evidence for this theme is mixed, comprising survey-based and program evaluation studies, with varying levels of methodological rigor, indicating moderate strength but limited consistency across different institutional contexts.

3.6. Theme 6: Research and Innovation-Based Leadership Learning

Research and innovation-based learning is increasingly recognised as a meaningful pathway for developing leadership capacity within higher education. Leadership learning in this context is closely linked to inquiry, problem solving, and active participation in knowledge creation. Undergraduate research is widely viewed as a powerful learning environment in which students develop leadership by engaging directly in investigation, collaboration, and intellectual responsibility (Healey & Jenkins, 2009). Research-based learning further supports leadership development by encouraging critical inquiry and participation in the construction of knowledge, positioning students as active contributors rather than passive recipients of information (Brew, 2013).

Inquiry-driven and problem-based pedagogies play a central role in fostering leadership through research and innovation. Collaborative problem solving and decision making are key features of problem-based learning, which has been shown to support leadership development by placing students in situations that require initiative, coordination, and shared responsibility (Prince & Felder, 2006). Inquiry-centred learning environments similarly promote leadership growth by requiring students to manage learning processes, work in teams, and engage with complex problems that mirror real-world challenges (Hmelo-Silver, 2004). Project- and problem-based learning further strengthen innovation-oriented leadership by situating students in authentic research contexts that demand creativity, accountability, and sustained collaboration (Kolmos et al., 2009).

The integration of research into academic curricula also contributes significantly to leadership development. Undergraduate research is widely recognised as a high-impact educational practice that enhances leadership learning by promoting deep engagement, independence, and responsibility for learning outcomes (Kuh, 2008). Student-as-researcher approaches extend this engagement by empowering students to take ownership of inquiry processes and participate actively in research communities, thereby strengthening leadership through responsibility and collaboration (Levy & Petrulis, 2012). Through these experiences, students move beyond knowledge consumption toward leadership through inquiry and innovation.

Research-led and innovation-oriented learning environments further support leadership development through collective knowledge production. Inquiry-based learning encourages leadership by engaging students in collaborative investigation, dialogue, and reflection, creating opportunities for shared leadership and critical engagement (Justice et al., 2009). Research-led teaching approaches similarly integrate innovation, teamwork, and reflective inquiry into learning processes, strengthening students'

leadership capacities through sustained and meaningful engagement (Walker et al., 2010). Embedding inquiry and innovation across curricula reinforces leadership development by enabling students to develop leadership capabilities through ongoing research engagement (Fung, 2017).

Research and innovation-based leadership learning thus extends leadership development into epistemic domains, positioning students as knowledge creators rather than passive recipients and reinforcing leadership as inquiry-driven practice. Overall, there is strong agreement in the literature that research and innovation-based learning provide an effective pathway for leadership development in higher education. Leadership emerges through inquiry, problem solving, and participation in knowledge creation, where students practise initiative, collaboration, and critical thinking. By integrating research and innovation into teaching and learning, higher education institutions create environments that support both independent and collective leadership learning, affirming the central role of research-based approaches in developing leadership capacity in contemporary higher education.

The evidence base for this theme is comparatively limited and often derived from conceptual or case-based studies, indicating that this area remains emergent and underexplored within the broader leadership development literature.

3.7. Integrative Synthesis of Leadership Development Pathways

Across the six themes, a clear pattern emerges whereby leadership development in higher education is most robust when multiple developmental pathways operate in alignment rather than isolation. Experiential engagement provides practical grounding, co-curricular spaces offer authentic leadership enactment, formal education contributes conceptual clarity, mentoring sustains identity development, professional skill formation strengthens employability, and research-based learning fosters critical inquiry and innovation. Together, these pathways form an interconnected ecosystem of leadership learning. The findings therefore support an integrated, student-centred framework in which leadership development is distributed across curricular, co-curricular, relational, and inquiry-based domains, rather than confined to a single pedagogical strategy.

While the synthesis identifies six coherent and interconnected pathways of leadership development, the strength of empirical evidence across these themes varies considerably. A substantial proportion of the reviewed studies are based on qualitative designs, single-institution case studies, or program evaluations, which, although rich in contextual insight, may limit generalisability. In particular, research on co-curricular leadership and mentoring is frequently associated with self-selection bias, as students who choose to participate in leadership activities may already possess higher levels of motivation or leadership inclination. As such, positive outcomes reported in these contexts should be interpreted with caution.

Furthermore, relatively few studies employ experimental or quasi-experimental designs capable of establishing causal relationships between leadership development interventions and student outcomes. This suggests that while there is strong conceptual and experiential support for integrated leadership development approaches, the robustness of causal evidence remains uneven across the literature.

In addition, contradictory or context-dependent findings are not extensively explored in existing studies. For instance, while co-curricular engagement is often associated with positive leadership outcomes, its effectiveness may vary depending on institutional support, programme quality, and student engagement levels. Similarly, the impact of formal leadership education may depend on the extent to which it is integrated with experiential and relational learning environments. These limitations indicate that the current evidence base, although conceptually coherent, would benefit from more rigorous, multi-institutional, and methodologically diverse research to strengthen the empirical foundations of integrated leadership development in higher education.

The proposed framework extends existing leadership development models by integrating multiple developmental pathways into a unified, student-centred ecosystem, offering a more holistic and practice-oriented approach compared to traditionally fragmented models.

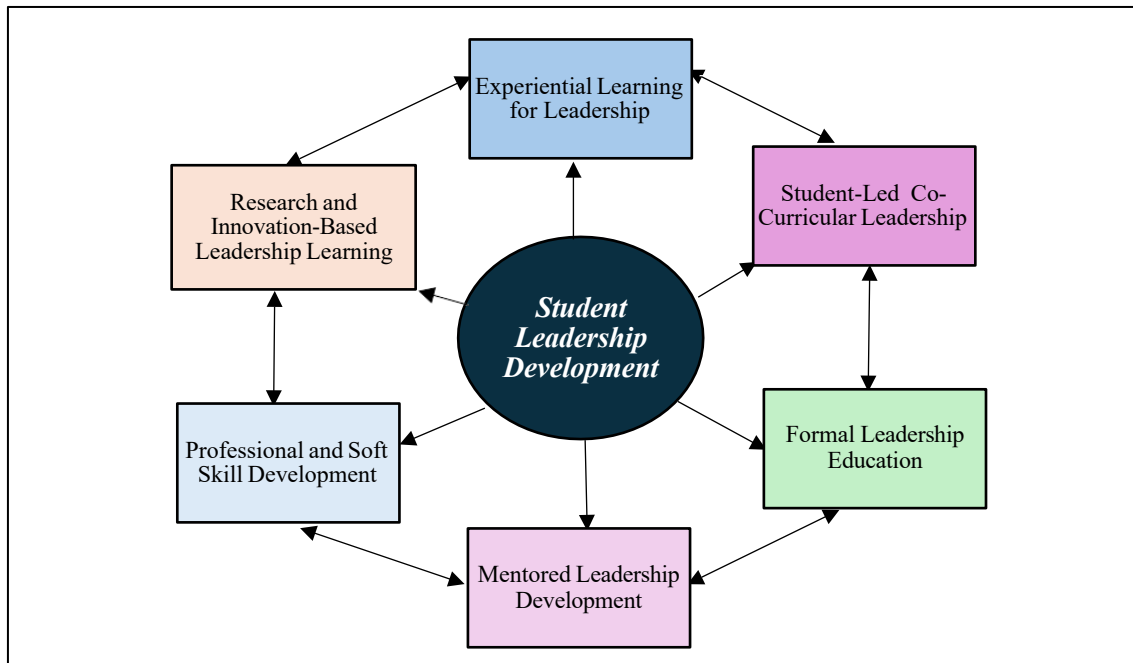


Figure 2. Integrated Student-Centred Leadership Development Framework

This figure illustrates the integrated student-centred leadership development framework, in which leadership development among university students emerges through the dynamic interaction of six interconnected pathways: experiential learning, student-led co-curricular engagement, formal leadership education, mentored leadership development, professional and soft skill development, and research and innovation-based learning.

At the centre of the framework is student leadership development, conceptualised as a holistic outcome shaped by the alignment of multiple learning environments. The surrounding pathways operate as an interconnected ecosystem, where each dimension reinforces and extends the others across curricular, co-curricular, relational, and inquiry-based domains. This model reflects the integrated nature of leadership learning, particularly within United States higher education contexts where these developmental pathways are institutionally embedded.

Table 3 summarises the six leadership development pathways identified in this review, along with their conceptual focus and supporting empirical evidence.

Table 3. Description of Integrated Student-Centred Leadership Development

Theme	Description	References	Evidence Type / Strength
Experiential Leadership Learning	• Leadership Develops Through Direct Experience • Students Learn by Practising Leadership in Real Contexts • Reflection Strengthens Leadership Understanding • Experience Helps Connect Theory And Practice	Astin & Astin (2000); Komives Et Al. (2005); Dugan & Komives (2010); Roberts (2012); Guthrie & Jenkins (2018); Eich (2008); Day Et Al. (2014); Priest & Gass (2018); Soria Et Al. (2015); Owen (2016); Kezar & Holcombe (2019); Mitchell Et Al. (2015)	Mixed (Predominantly Qualitative and Program-Based Studies) Limited Experimental Evidence
Student-Led Co-Curricular Leadership	• Leadership Develops Through Student-Led Activities and Initiatives • Students Practise Leadership by Organising and Leading Peers • Co-Curricular Spaces Offer Authentic, Informal Leadership Experiences • Engagement Outside The Classroom Supports Practical Leadership Development	Astin (1999); Dugan & Komives (2007); Komives Et Al. (2006); Kezar & Moriarty (2000); Soria Et Al. (2012); Haber (2011); Zimmerman-Oster & Burkhardt (1999); Owen Et Al. (2009); Luescher-Mamashela (2013); Dugan (2017)	Mostly Cross-Sectional and Correlational Studies Moderate Evidence
Formal Leadership Education	• Leadership Learning Is Supported Through Structured and Planned Instruction • Formal Courses Provide Systematic Leadership Knowledge and Framework • Curriculum-Based Programmes Ensure Intentional Leadership Development • Classroom Learning Strengthens Theoretical Understanding Of Leadership	Brungardt Et Al. (2006); Jenkins (2012); Rosch & Caza (2012); Komives Et Al. (2013); Priest & Thorpe (2017); Seemiller (2016); Guthrie, Jones, & Osteen (2017); Boatman Et Al. (2019); Day Et Al. (2014); Allen & Hartman (2008)	Include Quasi-Experimental and Curriculum-Based Evaluations Relatively Stronger Evidence
Mentored Leadership Development	• Leadership Develops Through Guided Support and Mentoring Relationships • Mentors Provide Feedback, Role Modelling, And Encouragement • Mentoring Strengthens Leadership Confidence and Identity	Crisp & Cruz (2009); Allen & Eby (2007); Komives Et Al. (2006); Dugan & Komives (2010); Packard (2016); Campbell Et Al. (2012); Gershenfeld (2014); Murphy & Johnson (2011); Johnson (2016); Dominguez & Hager (2013)	Largely Qualitative and Relational Studies Moderate Evidence

	• Guided Relationships Support Continuous Leadership Development		
Professional And Soft Skill Development	• Leadership Development Supports Professional and Soft Skill Growth • Students Build Communication, Teamwork, And Interpersonal Skills • Soft Skills Enhance Leadership Effectiveness and Employability • Leadership Learning Contributes To Transferable Professional Competencies	Seemiller (2013); Robles (2012); Heckman & Kautz (2012); Dugan (2017); Jackson (2014); Andrews & Higson (2008); Knight & Yorke (2004); Goleman (2000); Crosbie (2005); Laker & Powell (2011)	Mixed Evidence Including Survey-Based Studies Moderate To Strong
Research And Innovation-Based Leadership Learning	• Leadership Develops Through Research and Inquiry-Based Learning. • Students Practise Leadership by Solving Problems and Generating Knowledge. • Innovation-Oriented Activities Encourage Initiative and Critical Thinking. • Research Engagement Supports Independent and Collaborative Leadership Learning.	Healey & Jenkins (2009); Brew (2013); Prince & Felder (2006); Kolmos Et Al. (2009); Hmelo-Silver (2004); Fung (2017); Kuh (2008); Levy & Petrulis (2012); Justice Et Al. (2009); Walker Et Al. (2010)	Primarily Conceptual and Case - Based Studies Emerging Evidence

5. Limitations

This systematic literature review is subject to several limitations. First, the study was limited to publications indexed in the Scopus database. Although Scopus provides broad coverage of peer reviewed literature in social sciences and education, relevant studies indexed in other databases may not have been captured. Second, the search was restricted to English language publications, which may introduce language bias and limit the inclusion of scholarship published in other languages. Third, the review focused exclusively on studies conducted in the United States in order to ensure contextual consistency across higher education systems. While this strengthens comparability across institutional settings, it may limit the transferability of findings to other national or cultural contexts. In addition, although methodological quality was considered during the full text evaluation and thematic synthesis process, a formal risk of bias appraisal tool with numerical scoring was not applied. The synthesis therefore reflects interpretive assessment rather than structured quality grading. Future systematic reviews may incorporate formal appraisal instruments and extend the search across multiple databases and geographical contexts to further strengthen evidentiary breadth and methodological comparability.

6. Practical Implications and Implementation Roadmap

To translate the findings into actionable institutional strategies, Table 3 presents a prioritised implementation roadmap that categorises leadership development practices according to their implementation timeline, resource requirements, and potential institutional impact.

Table 4. Implementation Roadmap for Integrated Leadership Development

Timeline	Strategy	Resource Level	Expected Impact
Immediate	Embed leadership development within existing coursework and classroom activities	Low	Moderate
Medium-term	Expand structured mentoring and co-curricular leadership programmes	Medium	High
Long-term	Develop institution-wide integrated leadership development frameworks across multiple domains	High	Very high

This staged approach provides practical guidance for institutions to progressively implement integrated leadership development strategies based on available resources and institutional readiness.

7. Conclusion

This study synthesises leadership development in higher education into six interconnected pathways, highlighting that leadership development is most effective when experiential learning, co-curricular engagement, formal education, mentoring, professional skill development, and research-based learning are aligned within an integrated framework. The findings demonstrate that leadership development is inherently multidimensional and is best supported through a student-centred approach that spans curricular, co-curricular, relational, and inquiry-based domains.

Based on evidence derived from United States higher education contexts, an integrated approach to leadership development appears to be more effective. However, institutions are encouraged to adapt and evaluate these strategies within their own local contexts to ensure relevance and effectiveness across diverse educational environments. These findings carry important implications for higher education institutions, suggesting the need to design leadership development initiatives that are systematically integrated across multiple learning environments rather than implemented in isolation. Institutional stakeholders may consider embedding leadership development within both academic and co-curricular structures to enhance student engagement, leadership identity formation, and graduate employability.

Future research should prioritise more rigorous methodological designs, including longitudinal and experimental approaches, to strengthen causal understanding of leadership development outcomes. In addition, comparative studies across different national and institutional contexts would be valuable in assessing the transferability and applicability of integrated leadership development frameworks beyond the United States.

7. Ethics Statement

This study did not involve human participants, primary data collection, or the use of confidential data. The review was conducted using publicly available published literature. Therefore, ethical approval was not required.

8. Data Availability Statement

All data supporting the findings of this study are derived from publicly available peer reviewed journal articles indexed in the Scopus database. The list of included studies is available within the manuscript. No new datasets were generated or analysed beyond the published literature included in this review.

Generative AI Statement

During the preparation of this manuscript, AI-based tools were used to support specific aspects of the writing and analytical process. ChatGPT 4.0 assisted in refining language clarity and improving overall readability, while Elicit was used to help organise and interpret thematic patterns emerging from the reviewed studies. All outputs generated through these tools were carefully reviewed, critically evaluated, and substantively revised by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the published work.

References

- Allen, S. J., & Hartman, N. S. (2008). Leader development: An exploration of sources of learning. *Organization Development Journal*, 26(2), 75–87.
- Allen, T. D., & Eby, L. T. (2007). *The Blackwell handbook of mentoring: A multiple perspectives approach*. Blackwell Publishing.
- Andrews, J., & Higson, H. (2008). Graduate employability, soft skills versus hard business knowledge: A European study. *Higher Education in Europe*, 33(4), 411–422. <https://doi.org/10.1080/03797720802522627>
- Astin, A. W. (1999). Student involvement: A developmental theory for higher education. *Journal of College Student Development*, 40(5), 518–529.
- Astin, A. W., & Astin, H. S. (2000). Leadership reconsidered: Engaging higher education in social change. W. K. Kellogg Foundation.
- Boatman, N. J., Courtney, J. J., & Lee, W. Y. (2019). Leadership learning through academic courses: A review of leadership education practices in higher education. *Journal of Leadership Education*, 18(1), 1–15. <https://doi.org/10.12806/V18/I1/R1>
- Brew, A. (2013). Understanding the scope of undergraduate research: A framework for curricular and pedagogical decision making. *Higher Education*, 66(5), 603–618. <https://doi.org/10.1007/s10734-013-9626-x>
- Brungardt, C., Greenleaf, J., Brungardt, C., & Arensdorf, J. (2006). Majoring in leadership: A review of undergraduate leadership degree programs. *Journal of Leadership Education*, 5(1), 4–17. <https://doi.org/10.12806/V5/I1/R2>
- Campbell, C. M., Smith, M., Dugan, J. P., & Komives, S. R. (2012). Mentors and college student leadership outcomes: The importance of position and process. *Review of Higher Education*, 35(4), 595–625. <https://doi.org/10.1353/rhe.2012.0037>
- Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education*, 50(6), 525–545. <https://doi.org/10.1007/s11162-009-9130-2>
- Crosbie, R. (2005). Learning the soft skills of leadership. *Industrial and Commercial Training*, 37(1), 45–51. <https://doi.org/10.1108/00197850510699965>
- Day, D. V., Fleenor, J. W., Atwater, L. E., Sturm, R. E., & McKee, R. A. (2014). Advances in leader and leadership development: A review of 25 years of research and theory. *The Leadership Quarterly*, 25(1), 63–82. <https://doi.org/10.1016/j.leaqua.2013.11.004>
- Dominguez, N., & Hager, M. (2013). Mentoring frameworks: Synthesis and critique. *International Journal of Mentoring and Coaching in Education*, 2(3), 171–188. <https://doi.org/10.1108/IJMCE-03-2013-0014>
- Dugan, J. P. (2017). *Leadership theory: Cultivating critical perspectives*. Jossey-Bass.
- Dugan, J. P., & Komives, S. R. (2007). Developing leadership capacity in college students: Findings from a national study. *College Student Journal*, 41(2), 373–389.
- Dugan, J. P., & Komives, S. R. (2010). Influences on college students' capacities for socially responsible leadership. *Journal of College Student Development*, 51(5), 525–549. <https://doi.org/10.1353/csd.2010.0009>

- Eich, D. (2008). A grounded theory of high-quality leadership programs: Perspectives from student leadership programs in higher education. *Journal of Leadership & Organizational Studies*, 15(2), 176–187. <https://doi.org/10.1177/1548051808324099>
- Fung, D. (2017). *A connected curriculum for higher education*. UCL Press. <https://doi.org/10.14324/111.9781911576358>
- Gershenfeld, S. (2014). A review of undergraduate mentoring programs. *Review of Educational Research*, 84(3), 365–391. <https://doi.org/10.3102/0034654313520512>
- Goleman, D. (2000). Leadership that gets results. *Harvard Business Review*, 78(2), 78–90.
- Guthrie, K. L., & Jenkins, D. M. (2018). *The role of leadership educators: Transforming learning*. Information Age Publishing.
- Guthrie, K. L., Jones, T. B., & Osteen, L. (2017). Developing leadership educators: A descriptive study of leadership education graduates programs. *Journal of Leadership Education*, 16(2), 1–18. <https://doi.org/10.12806/V16/I2/R1>
- Haber, P. (2011). Peer education in student leadership programs: Responding to co-curricular challenges. *New Directions for Student Services*, 2011(133), 65–75. <https://doi.org/10.1002/ss.384>
- Healey, M., & Jenkins, A. (2009). *Developing undergraduate research and inquiry*. Higher Education Academy.
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451–464. <https://doi.org/10.1016/j.labeco.2012.05.014>
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235–266. <https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M. C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. McGill University.
- Jackson, D. (2014). Business graduate employability: Where are we going wrong? *Higher Education Research & Development*, 33(5), 905–918. <https://doi.org/10.1080/07294360.2013.806978>
- Jenkins, D. M. (2012). Examining the impact of leadership education on leadership capacity. *Journal of Leadership Education*, 11(2), 55–74. <https://doi.org/10.12806/V11/I2/RF4>
- Johnson, W. B. (2016). *On being a mentor: A guide for higher education faculty* (2nd ed.). Routledge.
- Justice, C., Rice, J., Roy, D., Hudspith, B., & Jenkins, H. (2009). Inquiry-based learning in higher education: Administrators’ perspectives on integrating inquiry pedagogy into the curriculum. *Higher Education*, 58(6), 841–855. <https://doi.org/10.1007/s10734-009-9228-3>
- Kezar, A., & Holcombe, E. (2019). *Shared leadership in higher education: Important lessons from research and practice*. American Council on Education.
- Kezar, A., & Moriarty, D. (2000). Expanding our understanding of student leadership development: A study exploring gender and ethnic identity. *Journal of College Student Development*, 41(1), 55–69.

Knight, P. T., & Yorke, M. (2004). *Learning, curriculum and employability in higher education*. RoutledgeFalmer.

Kolmos, A., de Graaff, E., & Du, X. (2009). Diversity of PBL: PBL learning principles and models. In X. Du, E. de Graaff, & A. Kolmos (Eds.), *Research on PBL practice in engineering education* (pp. 9–21). Sense Publishers.

Komives, S. R., Owen, J. E., Longerbeam, S. D., Mainella, F. C., & Osteen, L. (2005). Developing a leadership identity: A grounded theory. *Journal of College Student Development*, 46(6), 593–611. <https://doi.org/10.1353/csd.2005.0061>

Komives, S. R., Owen, J. E., Longerbeam, S. D., Mainella, F. C., & Osteen, L. (2006). A leadership identity development model: Applications from a grounded theory. *Journal of College Student Development*, 47(4), 401–418. <https://doi.org/10.1353/csd.2006.0048>

Komives, S. R., Wagner, W., & Associates. (2013). *Leadership for a better world: Understanding the social change model of leadership development* (2nd ed.). Jossey-Bass.

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Association of American Colleges and Universities.

Laker, D. R., & Powell, J. L. (2011). The differences between hard and soft skills and their relative impact on training transfer. *Human Resource Development Quarterly*, 22(1), 111–122. <https://doi.org/10.1002/hrdq.20063>

Levy, P., & Petrulis, R. (2012). How do first-year university students experience inquiry and research, and what are the implications for the practice of inquiry-based learning? *Studies in Higher Education*, 37(1), 85–101. <https://doi.org/10.1080/03075079.2010.499166>

Luescher-Mamashela, T. M. (2013). Student representation in university decision making: Good reasons, a new lens? *Studies in Higher Education*, 38(10), 1442–1456. <https://doi.org/10.1080/03075079.2011.625496>

Mitchell, T. D., Donahue, D. M., & Young-Law, C. (2015). Service learning as a pedagogy of whiteness. *Equity & Excellence in Education*, 48(4), 612–629. <https://doi.org/10.1080/10665684.2015.1060131>

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>

Murphy, S. E., & Johnson, S. K. (2011). The benefits of a long-lens approach to leader development: Understanding the seeds of leadership. *The Leadership Quarterly*, 22(3), 459–470. <https://doi.org/10.1016/j.leaqua.2011.04.004>

Owen, J. E. (2016). The leadership identity development model: Perspectives from college students. *Journal of College Student Development*, 57(1), 1–16. <https://doi.org/10.1353/csd.2016.0001>

Owen, J. E., Komives, S. R., & Longerbeam, S. D. (2009). Developing leadership identity: A grounded theory. *Journal of Leadership Education*, 8(1), 1–15. <https://doi.org/10.12806/V8/I1/RF1>

Packard, B. W. L. (2016). *Successful STEM mentoring initiatives for underrepresented students: A research-based guide for faculty and administrators*. Stylus Publishing.

- Priest, S., & Gass, M. A. (2018). *Effective leadership in adventure programming* (3rd ed.). Human Kinetics.
- Priest, S., & Thorpe, S. (2017). Experiential leadership training: Integrating theory and practice. *Journal of Experiential Education*, 40(4), 325–341. <https://doi.org/10.1177/1053825917714737>
- Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education*, 95(2), 123–138. <https://doi.org/10.1002/j.2168-9830.2006.tb00884.x>
- Roberts, D. C. (2012). *Leadership education: A critical analysis and review of the literature*. Jossey-Bass.
- Robles, M. M. (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. *Business Communication Quarterly*, 75(4), 453–465. <https://doi.org/10.1177/1080569912460400>
- Rosch, D. M., & Caza, A. (2012). Attitudes and beliefs about leadership: Implications for student leadership development. *Journal of College Student Development*, 53(3), 430–448. <https://doi.org/10.1353/csd.2012.0042>
- Saidi, Z. A. & Abd Aziz, M. F. (2025). Empowering student leadership in higher education: A literature review on key competencies for effective leadership. *Journal of Institutional Research South East Asia*, 23(1), 376-399.
- Saidi, Z.A., Asri, A.S., & Abd Aziz, M.F. (2025). Exploring the relationship between participation, needs, readiness, and planning in university student volunteering programs. *Journal of Institutional Research South East Asia*, 23(1), 401-417.
- Seemiller, C. (2013). *The student leadership competencies guidebook: Designing intentional leadership learning and development*. Jossey-Bass.
- Seemiller, C. (2016). *Student leadership competencies: A reference handbook*. Jossey-Bass.
- Soria, K. M., Johnson, M., & Mitchell, T. D. (2012). Experiential learning and leadership development through student involvement. *Journal of Leadership Education*, 11(2), 44–61. <https://doi.org/10.12806/V11/I2/RF3>
- Soria, K. M., Nobbe, J., & Fink, A. (2015). Examining the intersections between undergraduates' engagement in community service and development of socially responsible leadership. *Journal of Leadership Education*, 14(1), 1–16. <https://doi.org/10.12806/V14/I1/R1>
- Walker, M., Hall, J., & Ramsden, P. (2010). Engaging students in research and inquiry: A framework for higher education. *Studies in Higher Education*, 35(6), 697–710. <https://doi.org/10.1080/03075070903316745>
- Zimmerman-Oster, K., & Burkhardt, J. C. (1999). Leadership in the making: Impact and insights from leadership development programs in higher education. *Journal of Leadership Studies*, 6(3–4), 50–66. <https://doi.org/10.1177/107179199900600304>

The Relationship Between Assertiveness, Social Skills, And Psychological Well-Being Among Public University Students in Malaysia

ABSTRACT

Authors: Harizah Izyan Samsudin¹, Wan Norhayati Wan Othman^{1*}, Ahmad Fauzi Mohd Ayub¹ and Wan Marzuki Wan Jaafar¹

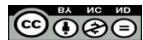
Affiliation: ¹Faculty of Educational Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

*Corresponding author:

wannorhayati@upm.edu.my

Received: 31/12/2025 |
First Revision: 14/1/2026 |
Second Revision: 15/4/2026 |
Accepted: 15/5/2026

This work is licensed under a



Creative Commons Attribution 4.0 International License

APA citation for this article:

This study examined the relationships between assertiveness, social skills, and psychological well-being, as well as the predictive contributions of assertiveness and social skills to psychological well-being among first-year undergraduate students in Malaysian public universities. A cross-sectional correlational research design was employed, and data were collected from 1,206 students using a multi-stage sampling procedure. Three established instruments were used, namely the Rathus Assertiveness Schedule (RAS), the Social Skills Inventory (SSI-Del-Prette), and Ryff's Psychological Well-Being Scale. The instruments demonstrated strong internal consistency, with Cronbach's alpha values of .93, .94, and .88, respectively. The findings indicated that assertiveness ($r = .259, p < .001$) and social skills ($r = .344, p < .001$) were significantly and positively associated with psychological well-being. In addition, multiple regression analysis showed that both assertiveness ($\beta = .192, p < .001$) and social skills ($\beta = .301, p < .001$) significantly predicted psychological well-being, explaining 15.3% of the variance. Hierarchical regression analysis further revealed that these relationships remained significant after controlling for demographic variables, which were not significant predictors. Overall, the findings suggest that interpersonal competencies, particularly social skills and assertiveness, are important factors associated with students' psychological well-being. These results have practical implications for higher education institutions, highlighting the importance of developing programmes that enhance students' interpersonal and social-emotional competencies to support their psychological well-being.

Keywords: Assertiveness, Social Skills, Psychological Well-Being, University Students, Regression Analysis, Malaysia

1. Introduction

University students experience a critical period of transition characterised by increased academic demands, social adjustment, and the development of personal identity. During this phase, students are required to adapt to new learning environments, establish interpersonal relationships, and manage expectations related to academic performance and future careers. These challenges may influence students' emotional functioning, coping abilities, and overall psychological well-being (Keyes, 2005; Ryff, 2014). Psychological well-being has become an important focus in higher education research, as it reflects positive functioning, life satisfaction, and the ability to manage stress effectively (Ryff, 1989). Students with higher psychological well-being tend to demonstrate greater resilience, emotional stability, and adaptive coping in response to academic and social pressures. However, beyond the examination of psychological distress, there is an increasing need to explore positive psychosocial competencies that are associated with students' well-being and overall functioning.

Among these competencies, assertiveness and social skills are considered important for interpersonal and emotional functioning. Assertiveness refers to the ability to express thoughts, needs, and emotions confidently and respectfully without violating the rights of others (Rathus, 1973). Students who are more assertive may be better able to communicate effectively and navigate social interactions. Similarly, social skills including coping and self-assertion with risk, self-assertion in expressing positive affect, conversation and social confidence, self-exposure to strangers and new situations, and self-control of aggressiveness are associated with social adjustment, peer acceptance, and psychological functioning (Del Prette & Del Prette, 2005). Previous research has reported that assertiveness and social skills are positively associated with interpersonal relationships, adaptive coping, and mental health outcomes among students (Habib & Ahmad, 2024; Ibrahim, 2023). In addition, psychological well-being has been linked to academic engagement, self-esteem, and emotional resilience (Keyes, 2005; Diener et al., 2018). However, most existing studies have focused on psychological distress rather than positive well-being, particularly within the Malaysian higher education context (Chelvarayan et al., 2023; Mohamad et al., 2024).

Furthermore, empirical evidence examining the combined associations of assertiveness and social skills with psychological well-being remains limited. Given this gap, there is a need to better understand how these psychosocial variables are related to psychological well-being among university students. Accordingly, this study aims (i) examine the relationships between assertiveness, social skills, and psychological well-being; (ii) determine whether assertiveness and social skills significantly predict psychological well-being; and (iii) examine whether these relationships remain significant after controlling for demographic variables, namely gender, field of study, and socioeconomic status, among students in Malaysian public universities. The findings are expected to provide empirical evidence to inform future research and support the development of student support initiatives in higher education settings.

2. Research Hypotheses

Based on the objectives of the study and the theoretical considerations discussed in the literature, a set of hypotheses was formulated to empirically examine the relationships and predictive contributions among the study variables. Given the cross-sectional design, the hypotheses are framed to test statistical relationships and predictive associations rather than causal effects. Specifically, the hypotheses address (i) the relationships between assertiveness, social skills, and psychological well-being, and (ii) the extent to which assertiveness and social skills contribute to the prediction of psychological well-being, both independently and after controlling for demographic variables.

1. H_{a1} : There is a significant relationship between assertiveness and psychological well-being.
2. H_{a2} : There is a significant relationship between social skills and psychological well-being.
3. H_{a3} : Assertiveness and social skills significantly predict psychological well-being.

4. H_a : Assertiveness and social skills significantly predict psychological well-being after controlling for demographic variables.

3. Literature Review

3.1. Assertiveness and Psychological Well-Being

Assertiveness is commonly defined as the ability to express one's thoughts, feelings, and needs confidently while respecting the rights of others, and it is considered a key psychosocial competency in interpersonal functioning (Rathus, 1973; Alberti & Emmons, 2017). In higher education contexts, assertiveness has been linked to healthier communication patterns, emotional regulation, and adaptive coping, which contribute to students' psychological well-being and adjustment (Eslami et al., 2018; Zhang et al., 2020). Psychological well-being refers to positive functioning across domains such as autonomy, self-acceptance, purpose in life, and environmental mastery (Ryff, 1989). Prior studies consistently report that students with higher assertiveness tend to experience higher self-esteem, greater resilience, and lower emotional distress, indicating that assertiveness may function as a personal resource that promotes well-being (Mesurado et al., 2019). Evidence from cross-sectional studies shows significant positive correlations between assertiveness and overall psychological well-being as well as related constructs such as empowerment, social adjustment, and life satisfaction (Plantade-Gipch et al., 2023; Rahayu et al., 2023).

Intervention-based studies further support this association. Assertiveness-training programmes have been shown to improve students' confidence, communication competence, and mental health outcomes, providing preliminary causal evidence for the beneficial role of assertiveness in enhancing psychological functioning (Ganji et al., 2022; Ilnytska & Onischenko, 2022). In Malaysian and regional student populations, assertiveness has also been associated with psychological empowerment and positive self-beliefs, particularly in professional and health-science education settings (Sabrina et al., 2022; Şen et al., 2022). However, findings also indicate that the relationship may vary depending on cultural norms, population characteristics, and the type of well-being construct measured. Studies suggest that assertiveness demonstrates stronger associations with empowerment and social adjustment outcomes, and comparatively smaller yet significant relationships with global well-being indicators (Rahayu et al., 2023; Plantade-Gipch et al., 2023). Despite these contextual variations, the literature generally supports the conclusion that assertiveness is positively linked to psychological well-being among university students.

3.2. Social Skills and Psychological Well-Being

Psychological well-being is a multidimensional construct encompassing individuals' positive functioning, emotional balance, self-acceptance, autonomy, and quality of interpersonal relationships. Within higher education contexts, psychological well-being has been linked to students' academic engagement, resilience, and overall adjustment to university life (García-Álvarez et al., 2021; Tekin, 2021). Based on the literature, social skills may be broadly defined as learned interpersonal competencies that enable individuals to interact effectively, communicate appropriately, regulate behaviour, solve social problems, and maintain positive relationships with others (Betinelli et al., 2025; Elliott et al., 2015; Riggio, 2020). Although definitions vary across scholars, most agree that social skills involve both behavioural and cognitive elements and are essential for successful functioning in academic, professional, and everyday social contexts (Dowd & Tierney, 2022; Lent et al., 2019; Trower, 1980). A growing body of empirical research demonstrates that students with stronger social skills report higher levels of psychological well-being across diverse cultural and educational contexts (Rosli & Perveen, 2021; Arhuis-Inca & Ipanaqué-Zapata, 2023).

Evidence from multiple cross-sectional studies indicates a positive relationship between various social skills constructs and psychological well-being. For instance, emotional intelligence has been found to

correlate moderately with psychological well-being among university students (Tekin, 2021; Tahir & Jabeen, 2022). Similarly, social connectedness has been shown to positively predict well-being indicators such as happiness, vitality, and flourishing (Yelpaze et al., 2021). Studies focusing on prosocial behavior also report positive associations with psychological well-being, particularly with dimensions related to positive relations with others (Rosli & Perveen, 2021; Liaqat & Deeba, 2025). Moreover, several studies highlight that the relationship between social skills and psychological well-being is often mediated by emotional and cognitive factors. Emotional expression fully mediated the relationship between social problem-solving skills and psychological well-being among nursing students (Eroğlu & Öner, 2023), while social self-efficacy mediated the relationship between positive thinking and life satisfaction among first-year university students (Amin et al., 2021).

These findings suggest that social skills enhance well-being not only directly but also by strengthening students' confidence in their interpersonal abilities and emotional regulation capacities. Although most studies report moderate to strong associations, weaker correlations have also been documented when broad or generalized measures of social skills are used (Tekin, 2021). This variation suggests that psychological well-being is influenced by multiple interacting factors beyond social competence alone. Nevertheless, the consistency of positive findings across studies supports the theoretical assumption that social skills are an important psychosocial resource for students in higher education.

3.3. The Interrelationship between Assertiveness, Social Skills, and Psychological Well-Being

Psychological well-being among university students has become an increasing concern due to rising academic, social, and emotional demands. It is commonly conceptualised as involving positive functioning, emotional balance, self-acceptance, and the ability to maintain satisfying interpersonal relationships. Evidence from Malaysian public university students indicates notable levels of psychological distress, including depression, anxiety, and stress, highlighting the importance of examining psychosocial factors associated with students' well-being (Latiff et al., 2014; Rathakrishnan et al., 2022). Within this context, assertiveness and social skills may be understood as two related but distinct interpersonal competencies that are relevant to students' adjustment and well-being.

Assertiveness is often described as lying on a behavioural continuum between passivity and aggression, representing a balanced form of self-expression that allows individuals to communicate their views firmly without violating the rights of others (Alberti & Emmons, 2017; Lazarus & Folkman, 1984). Conceptually, assertiveness has evolved from being viewed as a personality characteristic to being recognised as a social and communication skill. Early work by Salter (1949) linked assertiveness to emotional expressiveness, while later scholars conceptualised it as a form of social competence. Lazarus (1973) identified several core assertive abilities, including expressing one's needs openly, saying "no," communicating both positive and negative feelings, and initiating, maintaining, and ending conversations appropriately. More recent definitions similarly emphasise that assertiveness involves direct, honest, and respectful communication that protects one's own rights without undermining those of others (Afdal et al., 2019; Berger et al., 2019; Comstock, 2020; Oducado, 2021; Scott, 2018; Shanmugam & Kathyayini, 2017; Speed et al., 2018).

Social skills, in contrast, represent a broader set of learned interpersonal behaviours that enable individuals to interact and relate harmoniously and effectively with others, while also improving social acceptance and reinforcement in social contexts (Radley & Dart, 2022). These include communication, cooperation, empathy, emotional regulation, and conflict management (Del Prette & Del Prette, 2005). Social skills have been linked to academic success, professional readiness, emotional intelligence, and successful adaptation to the workplace, indicating that they are a core component of students' broader developmental functioning (Betinelli et al., 2025; Gul et al., 2023; Shengyao et al., 2024; Suleman et al., 2022; Vanhove et al.,

In the Malaysian context, the importance of social skills extends beyond interpersonal functioning to include employability and graduate readiness. Despite national development initiatives and increasing job opportunities, graduate unemployment remains a persistent issue, suggesting a mismatch between labour market expectations and graduates' competencies (Moo & Wan, 2023). Employers increasingly emphasise soft skills such as communication, teamwork, adaptability, and interpersonal effectiveness in addition to technical knowledge (Khazanah Research Institute, 2018;2020; World Economic Forum, 2025). These patterns suggest that social skills are relevant not only to career development but also to students' broader adjustment and well-being during university life. At the student level, evidence suggests that weaker social skills are associated with higher social anxiety, reduced interpersonal confidence, and difficulties in academic and social adjustment (Kipli et al., 2022). This is especially relevant for first-year university students, who must navigate new social environments, establish peer relationships, and adapt to academic demands simultaneously. It is also important to distinguish social skills from social support.

Cross-cultural evidence further supports the interrelationship among these constructs. Demir et al. (2012) demonstrated that social skills are associated with happiness through friendship quality, suggesting that interpersonal competence enhances well-being through the development of meaningful social relationships. Similarly, other studies indicate that social skills are associated with academic adjustment, resilience, and reduced emotional distress (Soares et al., 2021) while assertiveness is associated with effective self-expression and interpersonal confidence (Sial et al., 2021). Taken together, these findings suggest that students who are able to communicate effectively, express themselves appropriately, and maintain constructive social relationships may be more likely to experience positive psychological functioning.

Despite these findings, empirical studies that examine assertiveness, social skills, and psychological well-being simultaneously remain limited, particularly in the Malaysian higher education context. Most local research has focused on general soft skills or psychological distress, rather than examining how specific interpersonal competencies relate to positive indicators of well-being. This gap is important because first-year university students are undergoing a critical transition period in which both assertiveness and social skills may play a key role in psychological adjustment. Overall, the literature suggests that assertiveness and social skills are positively associated with psychological well-being, although the strength and expression of these relationships may vary across cultural and contextual settings. These interpersonal competencies may therefore be considered important psychosocial resources associated with students' well-being, particularly during periods of transition, uncertainty, and heightened stress, such as the post-pandemic university environment.

3.4. Demographic Factors and Study Variables

Previous research indicates that demographic factors such as gender, field of study, and socioeconomic status (SES) may be associated with variations in social skills and psychological well-being among university students. However, the findings across studies remain inconsistent, suggesting that these factors may function as background influences rather than primary determinants. Studies examining gender differences in social skills report mixed findings. For example, Carneiro Machado et al. (2020) found that gender did not significantly influence overall social skills; however, female students demonstrated higher scores in specific situations such as assertiveness in expressing positive emotions. Similarly, Prada et al. (2022) reported that female students exhibited stronger teamwork-related social skills, although differences were not consistent across all dimensions.

In relation to psychological well-being, several studies indicate no significant overall gender differences (Salleh & Mustaffa, 2016; Pumpuang et al., 2021). Nevertheless, some evidence suggests that females may score higher in dimensions such as positive relationships and purpose in life, whereas males may demonstrate higher autonomy (Iqbal & Nishat, 2017; Martud et al., 2019). These variations are often explained by socialisation processes, where males are encouraged towards independence and females towards relational connectedness. The field of study has also been linked to differences in both social

skills and psychological well-being. Research shows that students in programmes with higher practical and interpersonal engagement, such as nursing, tend to demonstrate stronger social skills compared to students in more theoretical disciplines (Carneiro Machado et al., 2020).

Similarly, studies on psychological well-being indicate that students in fields such as engineering and health sciences may report lower levels of well-being due to heavier academic workloads and practical demands (Panahi et al., 2013; Aisyah et al., 2021). In contrast, students in social science-related programmes may experience relatively better emotional adjustment due to differences in learning environments and academic expectations. Socioeconomic status (SES) is another factor that has been associated with students' social and psychological functioning. Students from lower SES backgrounds may experience challenges in forming and maintaining social relationships due to limited access to social resources and support systems (Gul et al., 2023). In contrast, students from higher SES backgrounds tend to benefit from stronger family and community support, which facilitates social engagement and interpersonal confidence.

In terms of psychological well-being, research indicates that students from higher SES backgrounds generally report better outcomes across multiple dimensions, including environmental mastery, self-acceptance, and life purpose (Dougall et al., 2023; Rubin, 2020). However, subjective SES has been found to be a stronger predictor of psychological functioning compared to objective SES indicators (Navarro-Carrillo et al., 2020). Overall, the literature demonstrates that gender, field of study, and SES may be associated with variations in social skills and psychological well-being. However, the findings remain inconsistent across contexts and populations. Moreover, most studies have examined these factors independently rather than considering them alongside key psychosocial competencies such as assertiveness and social skills. Within the Malaysian higher education context, empirical research examining these demographic differences remains limited, particularly among first-year undergraduate students. Therefore, these demographic variables are included in the present study as control variables to account for potential background influences when examining the associations between assertiveness, social skills, and psychological well-being.

3.5. Conceptual Framework of the Study

Based on the literature reviewed, this study proposes a conceptual framework in which assertiveness and social skills are treated as independent variables associated with psychological well-being as the dependent variable. Demographic variables, namely gender, field of study, and socioeconomic status, are included as control variables to account for potential variations in students' backgrounds. The framework assumes that higher levels of assertiveness and social skills are associated with higher levels of psychological well-being. However, consistent with the cross-sectional design of the study, these relationships are interpreted as statistical associations and predictive contributions rather than causal effects.

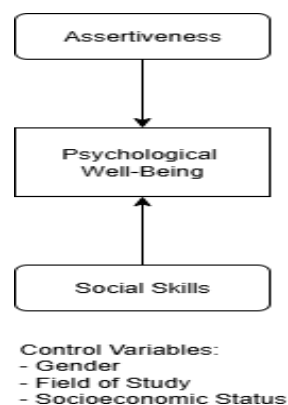


Figure 1. Conceptual Framework of Study

4. Methodology

4.1. Research Design

This study employed a quantitative correlational research design using a cross-sectional approach to examine the relationships between assertiveness, social skills, and psychological well-being among first-year undergraduate students in Malaysian public universities. A correlational design was considered appropriate as the study aimed to identify the strength and direction of relationships among variables without manipulating them. The cross-sectional approach enabled data to be collected at a single point in time, providing a snapshot of students' psychosocial competencies and well-being.

4.2. Population and Sampling

The target population of this study comprised first-year undergraduate students enrolled in Malaysian public universities during the 2021/2022 academic session. First-year students were selected as they are in a transitional phase of higher education, involving significant academic, social, and psychological adjustment. A multi-stage sampling procedure was employed to ensure representativeness across geographical regions and institutional contexts. In the first stage, cluster random sampling was applied due to the large and geographically dispersed population across public universities in Malaysia. These universities were grouped into five zones based on the classification by the Ministry of Higher Education, namely Central, Southern, Northern, Eastern, and Borneo. One university was randomly selected from each zone using a simple random technique marble-from-jar method, resulting in the selection of five universities: Universiti Teknologi MARA (UiTM), Universiti Malaysia Pahang (UMP), Universiti Teknikal Malaysia Melaka (UTeM), Universiti Sains Malaysia (USM), and Universiti Malaysia Sabah (UMS). This approach ensured representation across different geographical regions and institutional categories.

Based on institutional data, the total population of first-year undergraduate students across the five selected universities was 11,803 students. According to Krejcie and Morgan (1970), the minimum required sample size for a population of this size is 377 respondents, which is sufficient to achieve a 95% confidence level with a 5% margin of error. To account for potential incomplete, damaged, or unreturned questionnaires, an additional 10% was added (Mohd Najib, 2005), resulting in a proposed minimum sample size of 415 students. However, for this study, proportional stratified sampling was employed to determine the sample size required from each university. This technique is appropriate for populations that are not homogeneous (Creswell, 2005), as it ensures that each subgroup is adequately represented.

The selected universities differ in terms of institutional categories and geographical locations, with Universiti Sains Malaysia (USM) classified as a research university, Universiti Teknologi MARA (UiTM) and Universiti Malaysia Sabah (UMS) as comprehensive universities, and Universiti Teknikal Malaysia Melaka (UTeM) and Universiti Malaysia Pahang (UMP) as focused universities. Therefore, proportional representation was necessary to avoid underrepresentation of specific groups. In this study, approximately 10% of the population from each university was selected, resulting in a target sample size of 1,180 students (UiTM = 275, UTeM = 209, USM = 236, UMP = 207, UMS = 253). This sample size exceeds the minimum requirement and is considered adequate to represent the population (Fraenkel & Wallen, 1993). In the third stage, systematic random sampling was used to select respondents. Student lists were arranged alphabetically, and every *n*th student was selected until the required sample size for each university was achieved. This method ensured equal probability of selection and reduced sampling bias.

However, the sample for this study was drawn exclusively from the five selected universities which are UiTM, UMP, UTeM, USM, and UMS. While these universities represent multiple geographical zones in Malaysia and provide a diverse sample, the findings of this study are limited to the selected

institutions and should be interpreted with caution, as they may not be fully generalisable to all Malaysian public universities. For data collection, questionnaires were administered using an online platform (Google Forms). A total of 1,330 questionnaires were distributed across the five selected universities. Of these, 1,206 completed responses were received and retained for analysis, while 124 questionnaires were not returned, resulting in an overall response rate of 90.7%.

At the institutional level, slight variations in response rates were observed. Universiti Teknologi MARA (UiTM) recorded 275 responses from 310 distributed questionnaires (88.7%), Universiti Teknikal Malaysia Melaka (UTeM) recorded 209 responses from 236 distributed questionnaires (88.6%), Universiti Sains Malaysia (USM) recorded 236 responses from 266 distributed questionnaires (88.7%), Universiti Malaysia Pahang (UMP) recorded 207 responses from 233 distributed questionnaires (88.8%), and Universiti Malaysia Sabah (UMS) recorded 279 responses from 285 distributed questionnaires (97.9%), reflecting the highest response rate among the selected institutions.

Table 1. Sampling and Response Rate by University

University	Population (N)	Target Sample (≈10%)	Questionnaires Distributed	Responses Received	Non-response	Response Rate (%)
UiTM	2,750	275	310	275	35	88.7
UTeM	2,095	209	236	209	27	88.6
USM	2,360	236	266	236	30	88.7
UMP	2,068	207	233	207	26	88.8
UMS	2,530	253	285	279	6	97.9
Total	11,803	1,180	1,330	1,206	124	90.7

4.2.1 Ethical Considerations

This study adhered to standard ethical guidelines for research involving human participants. Prior to data collection, informed consent was obtained from all participants through an online Google Form. Participants were informed about the purpose of the study, the voluntary nature of their participation, and their right to withdraw at any time without penalty. Permission to conduct the study and collect data was obtained from the relevant institution. No personally identifiable information was collected, and all responses were recorded anonymously. The data were stored securely and used solely for academic research purposes. Formal institutional ethics approval was not obtained for this study, as it was conducted in accordance with institutional practices for low-risk survey research. However, all procedures were designed to ensure confidentiality, anonymity, and the protection of participants' rights.

4.2. Research Instrument

Research instruments are tools specifically designed for data collection. To produce a comprehensive and credible study, researchers may employ one or more instruments depending on the nature and purpose of the investigation (Fraenkel et al., 2018; Merriam & Tisdell, 2015). In the present study, a questionnaire was used as the main instrument to obtain the data required to answer the research questions and test the study hypotheses. According to Fraenkel et al. (2012), questionnaires may either be developed by the researcher or adapted from existing instruments in order to save time and ensure measurement consistency. Accordingly, the questionnaire used in this study was developed using an adapt-and-adopt approach.

The questionnaire comprised four main sections. Section A collected respondents' demographic information, including gender, ethnicity, field of study, place of residence, and socioeconomic status. Section B measured assertiveness using. Section C assessed students' social skills. Finally, Section D measured psychological well-being. The instruments used in this questionnaire had undergone expert validation in previous studies. However, the SSI-Del-Prette was subjected to content validation again

before being used in the present study. After the expert validation process was completed, the SSI was refined and subsequently tested for reliability through a pilot study together with the RAS and Ryff's Psychological Well-Being Scale.

4.3.1. Rathus Assertiveness Schedule (RAS)

Assertiveness in this study was measured using the Rathus Assertiveness Schedule (RAS), originally developed by Spencer Rathus (1973, 1978). The instrument was designed to assess the level of assertiveness among college students and adults. In the present study, the Malay version of the RAS was used. This version was adapted from previous local studies conducted by Abdullah (2005), Hawa (2014), Zainol Abidin (2016), Bakri (2016) and Jusoh et al.(2023). The RAS used in this study contained 30 items covering five dimensions of assertive behaviour, namely the right to mutual respect, the right to mutual feelings, the right to make mistakes, the right to say "no," and the right to ask questions and communicate. Fourteen items measured assertive behaviour, whereas sixteen items measured non-assertive behaviour. Responses were recorded on a five-point Likert scale ranging from 1 = strongly untrue to 5 = strongly true. Therefore, the minimum possible total score was 30 and the maximum possible total score was 150..As shown in Table 2, the instrument includes both positively and negatively worded items. All negatively items were reverse scored prior to analysis to ensure consistency in interpretation. Although the instrument comprises multiple dimensions, subscale scores were retained for descriptive purposes only and were not included in inferential analyses. The use of a composite mean score provides a parsimonious representation of assertiveness in regression analysis.

Table 2. Distribution of Positive and Negative Items for the Rathus Assertiveness Schedule (RAS)

Subscale	Description	Positive Items	Negative Items
1	Right to mutual respect	18, 20, 29	1, 2
2	Right to mutual feelings	8, 21	11, 15, 17, 24, 26
3	Right to make mistakes	22, 27, 28	4, 16, 19
4	Right to say "no"	3, 25	5, 9, 14, 23
5	Right to ask and communicate	6, 7, 10, 13	12, 30

4.3.2. Social Skills Inventory Del-Prette (SSI Del-Prette)

Social skills were measured using the Social Skills Inventory (SSI–Del-Prette), developed by Del Prette and Del Prette (2001, 2011, 2019). In this study, the Malay version adapted from Nor Azima (2008) was used. The SSI–Del-Prette has been widely used in Western contexts and has demonstrated satisfactory validity and reliability among college and university students aged between 18 and 25 years (Del Prette & Del Prette, 2019). In addition, the instrument has been used across clinical and non-clinical populations and in various fields such as health, counselling, education, forensic contexts, workplace settings, and higher education.

The version used in this study consisted of 38 items distributed across five factors, namely coping and self-assertion with risk, self-assertion in expressing positive affect, conversation and social confidence, self-exposure to strangers and new situations, and self-control of aggressiveness. Respondents were required to indicate their agreement using a four-point Likert scale ranging from 1 = never to 4 = always. The minimum possible total score was 38, while the maximum possible total score was 152. As shown in Table 3. The instrument includes both positively and negatively worded items. All negatively items were reverse scored prior to analysis to ensure consistency in interpretation. For the purpose of analysis, composite scores were computed by calculating the mean of all items to represent overall social skills. Subscale scores for each factor were also computed as mean values but were used for descriptive purposes only. The use of an overall composite score is consistent with prior studies and provides a parsimonious representation of psychological well-being in regression analysis.

Table 3. Distribution of Positive and Negative Items for the Social Skills Inventory (SSI Del-Prette)

Subscale	Description	Positive Items	Negative Items
1	Coping and self-assertion with risk	4, 5, 11, 21, 22, 29	25, 33, 34, 37
2	Self-assertion in expressing positive affect	3, 6, 10, 20, 28, 30, 35	13
3	Conversation and social confidence	12, 14, 17, 32	8, 19, 24
4	Self-exposure to strangers and new situations	1, 7, 31	9, 23, 26, 36
5	Self-control of aggressiveness	15, 16, 27, 38	2, 18

4.3.3. Ryff's Psychological Well-Being Scale

Psychological well-being was measured using Ryff's Psychological Well-Being Scale (Ryff, 1989), which comprises six dimensions, namely autonomy, environmental mastery, personal growth, purpose in life, positive relations with others, and self-acceptance. Ryff developed several versions of this instrument, including long, medium, and short forms. In the present study, the 42-item short version was used, with seven items representing each dimension. The Malay version used in this study was adapted from previous local studies conducted by Hasimah (2010), Norsayyidatina (2017), Roslan et al. (2017), and Salina and Rahimi (2015).

Responses were measured using a six-point Likert scale ranging from 1 = strongly disagree to 6 = strongly agree. Negative items were reverse scored before analysis. The minimum possible total score for the 42 items was 42, while the maximum possible score was 252. The 42-item version of Ryff's Psychological Well-Being Scale comprises six dimensions, as presented in Table 4. The instrument includes both positively and negatively worded items. All negatively worded items were reverse scored prior to analysis to ensure consistency in interpretation, where higher scores indicate higher levels of psychological well-being. For the purpose of analysis, a composite score was computed by calculating the mean of all items to represent overall psychological well-being. Mean scores for each dimension were also computed but were used for descriptive purposes only. The use of an overall composite score is consistent with prior studies and provides a parsimonious representation of psychological well-being in regression analysis.

Table 4. Distribution of Positive and Negative Items for Ryff's Psychological Well-Being Scale

Dimension	Description	Positive Items	Negative Items
1	Autonomy	1, 7, 25, 37	13, 19, 31
2	Environmental mastery	2, 8, 20, 38	14, 26, 32
3	Personal growth	9, 21, 33	3, 15, 27, 39
4	Positive relations with others	4, 22, 28, 40	10, 16, 34
5	Purpose in life	11, 29, 35	5, 17, 23, 41
6	Self-acceptance	6, 12, 24, 42	18, 30, 36

4.4. Instrument Validity

Instrument validity is crucial in ensuring that a measurement accurately reflects what it intends to measure (Azizi Yahya et al., 2018; Field, 2018). It is essential that the validation process is performed carefully to assess whether the instrument effectively measures the intended constructs. Validity can be classified into several types, including face validity, content validity, criterion validity, and construct validity (Taherdoost, 2016). In this study, the primary focus was on content validity, which evaluates whether each item in the instrument is suitable for its intended purpose.

For this research, the instruments used were all in Malay and adapted from previous studies. The Rathus Assertiveness Schedule (RAS) and Ryff's Psychological Well-Being Scale have been widely employed in local studies and have demonstrated strong validity. For example, Abdullah (2005) used the RAS to assess assertiveness in employees at the Johor State Customs Department, achieving a content validity of 0.73. Similarly, Hawa (2014) reported a content validity of 0.82 for the RAS in exploring assertiveness and organizational commitment.

Furthermore, Salina & Rahimi (2015) conducted a study testing the validity of Ryff's Well-Being Scale, obtaining a high content validity index (CVI) of 0.94. In contrast, Hasimah (2010) also found a high content validity score of 0.92 when using Ryff's scale to measure psychological well-being among adolescents. Since these two instruments demonstrated strong validity, the researcher focused on validating the Social Skills Inventory Del-Prette (SSI Del-Prette) (Del Prette & Del Prette, 2001; Nor Azima, 2008).

4.4.1 Process of Determining Content Validity for SSI Del-Prette

The Content Validity Index (CVI) was chosen to assess the content validity of the SSI Del-Prette instrument due to its advantages, such as being more cost-effective and practical than other methods (Tojib & Sugianto, 2006). The method involves obtaining feedback from a panel of experts to ensure detailed evaluation of each item, which provides a more comprehensive understanding than other methods like Cohen's Kappa or Tinsley-Weiss T Index. The CVI can be calculated through two methods which are the Item Content Validation Index (I-CVI) and the Sum of Content Validation Index (S-CVI). The I-CVI guides experts in evaluating and possibly revising items based on their relevance (Polit & Beck, 2006; Salbiah @ Salleh et al., 2017). In this study, external reviewers including experts in psychometrics and counseling, were selected to assess the SSI instrument. Expert selection was based on criteria such as experience, academic qualifications, and expertise in relevant fields, with six experts chosen for their qualifications and experience in evaluating instruments.

The expert evaluation used a 10-point Likert scale, where scores of 1 to 5 indicated items were unsuitable, and scores of 6 to 10 indicated items were suitable. Experts were also asked to provide qualitative feedback to improve the instrument's relevance to Malaysia's language, society, and culture. To calculate content validity, the study used the Content Validity Ratio (CVR) formula, which measures the agreement among experts on the relevance of each item (Lawshe, 1975). A CVR score of above 0.80 is recommended for new instruments, ensuring high content validity (Lau et al., 2018). In this study, the SSI instrument achieved an S-CVI of 0.99, indicating very high content validity. This confirms that the instrument's items are appropriate and accurately represent the constructs being measured.

4.5. Instrument Reliability

Reliability is essential for ensuring the stability, predictability, and accuracy of a research tool's findings (Hamidah Yusof et al., 2015). It refers to the consistency of results when the same instrument is used repeatedly (Azizi Ahmad, 2010). Several methods can assess reliability, including Kuder Richardson, Internal Consistency, Test-Retest, Inter-Rater Reliability, and Split-Half, with Cronbach's Alpha being the most commonly used (George & Mallery, 2016). To assess the reliability of the instruments in this study, the researcher used Cronbach's Alpha, a coefficient that measures internal consistency among the items in an instrument (Sidek Mohd Noh, 1998).

A Cronbach's Alpha value above 0.6 is considered acceptable, while values above 0.7 are ideal (Cresswell, 2016; Pallant, 2010). For exploratory studies, values as low as 0.6 may still be acceptable (Mohd Majid, 2005). Since the instruments used in the study had already been tested in previous research, they were re-tested within the context of Malaysian society, considering modifications such as phrasing and cultural context. A pilot study was conducted with 137 first-year undergraduate students from a university on the east coast of Malaysia, using the RAS, SSI Del-Prette, and Ryff's Psychological Well-Being Scales in Malay. The reliability of the instruments was assessed using Cronbach's Alpha to ensure their consistency and suitability for the research.

The results of the reliability tests showed that the RAS achieved a Cronbach's Alpha of 0.93, indicating excellent reliability for measuring assertiveness. The SSI Del-Prette instrument displayed an even higher Cronbach's Alpha of 0.94, confirming its very high reliability for assessing social skills. The Ryff's Psychological Well-Being Scale also demonstrated strong reliability with a Cronbach's Alpha of

0.88 which is considered high for evaluating psychological well-being. These results confirm that all three instruments exhibit strong internal consistency, making them reliable tools for measuring assertiveness, social skills, and psychological well-being in university students.

4.6. Normality of data

The normality of the data was assessed using both statistical indicators (skewness and kurtosis) and graphical methods (histogram and scatterplot of standardized residuals). Based on the descriptive statistics presented in Table 5, the skewness and kurtosis values for all variables were within the acceptable range of ± 1 , indicating that the data were approximately normally distributed. Specifically, assertiveness (RAS) showed a skewness value of 0.159 and kurtosis of 0.092, social skills (SSI) showed a skewness of -0.104 and kurtosis of 0.098, and psychological well-being (Ryff) recorded a skewness of 0.018 and kurtosis of -0.117. These values suggest that the distributions for all variables did not deviate substantially from normality.

In addition, graphical inspection of the histogram of standardized residuals demonstrated a bell-shaped distribution, further supporting the assumption of normality. The distribution of residuals appeared symmetrical and centred around zero, indicating no serious deviation from normal distribution. The scatterplot of standardized residuals also showed a random distribution of data points with no clear pattern or funnel shape. This suggests that the assumptions of normality and homoscedasticity were satisfied. Overall, the results indicate that the data met the normality assumption required for parametric statistical analyses, including correlation and regression. Furthermore, the large sample size ($N = 1,206$) supports the robustness of parametric tests, as the Central Limit Theorem suggests that the sampling distribution approximates normality. All inferential analyses were conducted using mean scores, and z-score standardisation was not applied.

Table 5. Descriptive Statistics for Study Variables

Variable	Mean	SD	Skewness	Kurtosis
Assertiveness (RAS)	2.792	0.390	0.159	0.092
Social Skills (SSI)	2.849	0.524	-0.104	0.098
Psychological Well-Being (Ryff)	3.202	0.548	0.018	-0.117

5. Findings and Discussion

This section presents the findings of the study in accordance with the research objectives. The results are organised into four main parts: (i) descriptive analysis of the study variables, (ii) examination of the relationships between assertiveness, social skills, and psychological well-being, (iii) analysis of the predictive contributions of assertiveness and social skills to psychological well-being, and (iv) hierarchical regression analysis to determine whether these relationships remain significant after controlling for demographic variables.

Descriptive statistics were first examined to provide an overview of the distribution of the variables. Subsequently, Pearson correlation analysis was conducted to assess the relationships among the variables. Multiple regression analysis was then performed to determine the extent to which assertiveness and social skills predict psychological well-being. Finally, hierarchical regression analysis was conducted to evaluate the predictive contributions of these variables after accounting for demographic factors, namely gender, field of study, and socioeconomic status. All analyses were conducted using appropriate parametric statistical techniques, as the assumptions of normality were met.

5.1. Descriptive Analysis of the Study Variables

Descriptive statistics were conducted to examine the mean scores, standard deviations, skewness, and kurtosis of the study variables. The results are presented in Table 5. The findings show that the mean score for assertiveness was 2.79 (SD = 0.39), while social skills recorded a mean score of 2.85 (SD = 0.52). Psychological well-being showed a higher mean score of 3.20 (SD = 0.55) compared to the other variables. These results indicate that, on average, students reported relatively higher psychological well-being compared to assertiveness and social skills. In addition, the skewness and kurtosis values for all variables were within the acceptable range of ± 1 , suggesting that the data were approximately normally distributed. Therefore, the assumptions for parametric statistical analyses were satisfied.

5.2. Examination of the Relationships Between Assertiveness, Social Skills, and Psychological Well-Being

Pearson correlation analysis was conducted to examine the relationships among the study variables. In this study, two alternative hypotheses were tested. The first hypothesis (H_{a1}) proposes that there is a significant relationship between assertiveness and psychological well-being, while the second hypothesis (H_{a2}) proposes that there is a significant relationship between social skills and psychological well-being. The results are presented in Table 6.

The findings indicated that assertiveness was significantly and positively correlated with psychological well-being, $r(1204) = .259$, $p < .001$. This result shows that the assumption of no relationship stated in H_{01} is not supported. Therefore, H_{a1} was supported, indicating that a significant relationship exists between assertiveness and psychological well-being.

Similarly, social skills were significantly and positively correlated with psychological well-being, $r(1204) = .344$, $p < .001$. Thus, H_{a2} was supported, indicating that a significant relationship exists between social skills and psychological well-being. In addition, assertiveness was significantly correlated with social skills, $r(1204) = .224$, $p < .001$, indicating that these variables are related constructs.

Table 6. Correlation Matrix of Study Variables (N=1206)

Variable	1	2	3
Assertiveness	—		
Social Skills	.224***	—	
Psychological Well-Being	.259***	.344***	—

Note. $N = 1,206$. ** $p < .001$.

The findings of this study revealed a significant positive relationship between assertiveness and psychological well-being among first-year university students. This result is consistent with previous research suggesting that assertiveness is associated with better emotional functioning and adjustment in higher education contexts. Prior studies have shown that assertiveness is positively related to academic, social, and personal-emotional adjustment, with social adjustment often demonstrating the strongest association. This indicates that assertive behaviour may facilitate more effective peer interaction and social integration within university environments (Rahayu et al., 2022).

From a psychosocial perspective, assertiveness can be understood as a personal resource that is associated with higher levels of confidence and perceived competence in managing academic and social demands. Empirical evidence indicates that students who demonstrate stronger assertive tendencies tend to report greater psychological empowerment and self-efficacy (Sabrina et al., 2021; Şen et al., 2022). In addition, previous studies have linked assertiveness with higher self-esteem, suggesting that individuals who are able to express themselves confidently are more likely to develop a positive sense of self-worth (Chandio et al., 2024; Chakraborty et al., 2020). Further support for this finding can be observed in studies examining broader indicators of well-being. Lower levels of assertiveness have been associated with reduced psychological and social well-being, while higher assertiveness has been linked to lower levels of stress across academic, social, and personal domains (Plantade-Gipch et al., 2023;

Hernández-Xumet et al., 2024). Taken together, these findings suggest that assertiveness is consistently associated with adaptive psychological outcomes.

However, it is important to note that the strength of the relationship observed in this study is modest. This suggests that while assertiveness is associated with psychological well-being, it is likely one of several contributing factors. Cultural considerations may also play a role, particularly in collectivist contexts such as Malaysia, where the expression of assertiveness may be moderated by social norms that emphasise harmony and interpersonal sensitivity. As such, the relationship between assertiveness and psychological well-being may vary depending on cultural expectations and social context. The present study also found a significant positive relationship between social skills and psychological well-being, which is consistent with previous research conducted in higher education settings (Tekin, 2021; Arhuis-Inca & Ipanaqué-Zapata, 2023). From a theoretical perspective, this finding aligns with psychosocial and positive psychology frameworks, which suggest that effective interpersonal functioning is associated with improved psychological well-being through the development of supportive relationships, emotional expression, and adaptive coping strategies (Eroğlu & Öner, 2023).

Students with stronger social skills are generally better able to communicate effectively, manage interpersonal conflicts, and maintain meaningful social connections. These competencies are associated with better emotional adjustment and overall mental health outcomes. In addition, prior research suggests that social skills may operate through mechanisms such as emotional regulation and social self-efficacy, further supporting their association with psychological well-being (Amin et al., 2021). The positive correlation observed between assertiveness and social skills in this study further suggests that these constructs are related aspects of interpersonal competence. Students who are more assertive may also be more capable of engaging in effective social behaviours, such as clear communication and appropriate self-expression. However, the relatively modest strength of this relationship indicates that assertiveness and social skills, while related, represent distinct constructs.

Overall, the findings of this study support existing literature indicating that both assertiveness and social skills are positively associated with psychological well-being among university students. However, consistent with the cross-sectional design of the study, these findings should be interpreted as statistical associations rather than causal relationships. From a practical perspective, the results suggest that higher education institutions may benefit from supporting the development of students' interpersonal competencies. Initiatives such as communication skills training, peer mentoring programmes, and social-emotional learning interventions may be beneficial in supporting students' adjustment and well-being. Future research is recommended to explore additional factors associated with psychological well-being, as well as to examine these relationships using longitudinal or experimental designs to better understand their direction and development over time.

5.3. Determine whether Assertiveness and Social Skills Predict Psychological Well-Being

This objective aimed to determine whether assertiveness and social skills predict psychological well-being. The alternative hypothesis for this objective (H_{a3}) proposes that assertiveness and social skills significantly predict psychological well-being. Multiple regression analysis was conducted, and the results are presented in Table 7. The regression model was statistically significant, $F(2, 1203) = 108.777$, $p < .001$, explaining 15.3% of the variance in psychological well-being ($R^2 = .153$). Both assertiveness ($\beta = .192$, $p < .001$) and social skills ($\beta = .301$, $p < .001$) were found to be significant predictors of psychological well-being. Therefore, H_{a3} was supported, indicating that assertiveness and social skills significantly predict psychological well-being.

Table 7. Analysis Predicting Psychological Well-Being (N=1206)

Predictor	B	SE	β	<i>t</i>	<i>p</i>
Assertiveness	0.269	0.038	0.192	7.026	< .001

Social Skills	0.315	0.029	0.301	11.018	< .001

Note. $N = 1,206$. $**p < .001$.

The findings of this study indicate that both assertiveness and social skills significantly predict psychological well-being among first-year university students. This suggests that these interpersonal competencies are important factors associated with variations in students' psychological well-being. The predictive relationships observed in this study are consistent with previous research, which highlights the role of interpersonal competencies in supporting students' psychological functioning. In particular, prior studies conducted in the Malaysian context have shown that social support often considered an outcome of effective social skills is associated with lower psychological distress and higher levels of well-being among university students (Abdullah et al., 2014; Rathakrishnan et al., 2022). These findings suggest that students who are able to communicate effectively and engage in positive social interactions may be more likely to develop supportive relationships that contribute to their emotional well-being.

From a theoretical perspective, the findings can be explained through interpersonal and psychosocial frameworks, which propose that effective social functioning enhances well-being through mechanisms such as social connectedness and emotional support. Cross-cultural research further indicates that social skills may contribute to psychological well-being indirectly through factors such as friendship quality and social connectedness (Demir et al., 2012). In this context, assertiveness and social skills may function in complementary ways, where assertiveness is associated with the ability to express personal needs and boundaries, while social skills are associated with maintaining and strengthening interpersonal relationships. Although both predictors were found to be significant, social skills demonstrated a relatively stronger contribution compared to assertiveness. This may suggest that, within the university context, the ability to maintain relationships and navigate social interactions plays a particularly important role in students' psychological well-being. The findings also carry practical implications for higher education institutions.

Given the reported prevalence of psychological distress among Malaysian university students (Latiff et al., 2014; Hamzah et al., 2023), programmes that support the development of interpersonal competencies may be beneficial. Interventions focusing on communication skills, peer interaction, and social-emotional development may contribute to improved student well-being. However, such initiatives should be implemented with consideration of cultural norms, particularly in collectivist contexts where interpersonal harmony is emphasised. Overall, the present findings provide empirical support that assertiveness and social skills are significant predictors of psychological well-being. While the amount of variance explained is modest, the results are statistically significant and theoretically meaningful, highlighting the importance of interpersonal competencies in understanding students' psychological well-being in higher education settings.

5.4. Examine whether Assertiveness and Social Skills Predict Psychological Well-Being after Controlling for Demographic Variables

This objective aimed to examine whether assertiveness and social skills predict psychological well-being after controlling for demographic variables. The alternative hypothesis for this objective (Ha₄) proposes that assertiveness and social skills significantly predict psychological well-being after controlling for demographic variables. A hierarchical multiple regression analysis was conducted, and the results are presented in Table 8. In Model 1, demographic variables (gender, field of study, and socioeconomic status) were entered as predictors. The results indicated that the model was not statistically significant, $F(3, 1202) = 0.584$, $p = .625$, explaining only 0.1% of the variance in psychological well-being ($R^2 = .001$). This finding suggests that demographic variables alone do not contribute significantly to explaining psychological well-being among the students.

In Model 2, assertiveness and social skills were added to the model. The model became statistically significant, $F(5, 1200) = 43.670$, $p < .001$, explaining 15.4% of the variance in psychological well-being

($R^2 = .154$). This indicates that the inclusion of assertiveness and social skills substantially improved the explanatory power of the model. Further analysis showed that both assertiveness ($\beta = .192, p < .001$) and social skills ($\beta = .301, p < .001$) remained significant predictors of psychological well-being, while demographic variables remained non-significant. Therefore, H_{a4} was supported, indicating that assertiveness and social skills significantly predict psychological well-being even after controlling for demographic variables.

Table 8. Hierarchical Regression Predicting Psychological Well-Being (N=1206)

Predictor	B	SE	B	t	p
Model 1					
Field of Study	-0.010	0.032	-0.009	-0.327	.744
Gender	0.014	0.032	0.013	0.438	.662
Socioeconomic Status	0.030	0.025	0.035	1.219	.223
Model 2					
Field of Study	-0.013	0.030	-0.012	-0.454	.650
Gender	0.026	0.030	0.024	0.897	.370
Socioeconomic Status	0.008	0.023	0.009	0.338	.735
Assertiveness	0.269	0.038	0.192	7.026	<.001
Social Skills	0.315	0.029	0.301	11.018	<.001

Note. $N = 1,206$. ** $p < .001$.

The findings of this study indicate that assertiveness and social skills remain significant predictors of psychological well-being even after controlling for demographic variables, namely gender, field of study, and socioeconomic status. This suggests that interpersonal competencies are more strongly associated with psychological well-being compared to demographic characteristics among first-year university students. The non-significant results observed for demographic variables in Model 1 indicate that gender, field of study, and socioeconomic status do not independently contribute to explaining variations in psychological well-being in this sample. This finding is consistent with previous research that has reported inconsistent or weak associations between demographic factors and psychological well-being among university students (Salleh & Mustaffa, 2016; Pumpuang et al., 2021). It also suggests that psychological well-being may be influenced more by individual and interpersonal factors rather than background characteristics alone.

At the same time, several studies point to contextual variation in these relationships. Gender has been identified as a potential moderating factor, with some studies reporting that female students tend to exhibit lower levels of assertiveness and higher levels of social anxiety compared to male students (Hayat et al., 2023; Bamoussa et al., 2023). However, other studies have found no significant gender differences, suggesting that the influence of gender may depend on cultural and situational factors (Nair et al., 2020). In addition, cultural and disciplinary contexts have been shown to shape the expression of assertiveness, particularly in collectivist societies and professional training environments where behavioural norms emphasise compliance and interpersonal harmony, such as in nursing education (Chakraborty et al., 2020; Sabrina et al., 2021). These mixed findings may help explain why demographic variables did not emerge as significant predictors in the present study.

The significant improvement in the model after the inclusion of assertiveness and social skills indicates that these variables provide meaningful explanatory value beyond demographic factors. This finding supports theoretical perspectives which emphasise the role of psychosocial competencies in shaping individuals' emotional functioning and overall well-being. In particular, interpersonal competence frameworks suggest that the ability to communicate effectively, express oneself appropriately, and maintain social relationships is associated with better psychological adjustment. Consistent with the findings in Objective 3, social skills demonstrated a relatively stronger contribution compared to assertiveness. This may indicate that, within the university context, the ability to establish and maintain social connections plays a particularly important role in students' psychological well-being.

Assertiveness, while still significant, may function as a complementary skill that supports effective communication within these relationships.

These findings also highlight the importance of considering interpersonal competencies as protective factors in promoting students' psychological well-being. Given that demographic variables did not significantly influence the outcome, interventions aimed at enhancing assertiveness and social skills may be broadly applicable across diverse student groups. However, such interventions should be implemented with consideration of cultural context, particularly in collectivist societies where interpersonal harmony and social sensitivity are emphasised. Overall, the findings suggest that assertiveness and social skills are significant predictors of psychological well-being beyond demographic influences. Although the variance explained by the model is modest, the results are statistically significant and theoretically meaningful, reinforcing the importance of interpersonal competencies in understanding students' psychological well-being in higher education settings.

6. Conclusion

This study examined the relationships and predictive contributions of assertiveness and social skills to psychological well-being among first-year university students in Malaysian public universities. The findings revealed that students reported moderate levels of assertiveness and social skills, alongside a relatively higher level of psychological well-being. The data were normally distributed, supporting the use of parametric analyses. The results further indicated that both assertiveness and social skills were significantly and positively associated with psychological well-being. In addition, both variables were found to be significant predictors of psychological well-being, with social skills demonstrating a relatively stronger contribution compared to assertiveness. Importantly, these relationships remained significant even after controlling for demographic variables, namely gender, field of study, and socioeconomic status, which were found to be non-significant predictors. Overall, the findings suggest that interpersonal competencies, particularly social skills and assertiveness, are important factors associated with psychological well-being among university students. Although the variance explained by the model was modest, the results are statistically significant and theoretically meaningful. These findings contribute to the growing body of literature highlighting the role of psychosocial competencies in supporting students' adjustment and well-being in higher education contexts.

7. Recommendation

Based on the findings, higher education institutions are encouraged to implement programmes that support the development of students' interpersonal competencies. Initiatives such as communication skills training, assertiveness development workshops, peer mentoring programmes, and social-emotional learning interventions may help enhance students' psychological well-being. In addition, university counselling services may integrate interpersonal skills training into existing mental health programmes to support students in managing academic and social challenges more effectively. Given that demographic variables were not significant predictors; such interventions may be applied broadly across diverse student groups. At the institutional level, universities may consider embedding interpersonal and psychosocial skill development into the curriculum, particularly for first-year students who are undergoing transitional adjustment. Structured programmes such as orientation modules, co-curricular activities, and student development courses may provide opportunities for students to enhance their communication and social competencies.

Future research is recommended to further leverage the strengths of large and diverse samples by examining whether the relationships among assertiveness, social skills, and psychological well-being differ across institutional and demographic subgroups. Specifically, moderation or multi-group analyses should be conducted to test whether these relationships vary by university context, gender, field of study, or socioeconomic status. Such analyses would provide a more nuanced understanding of whether

interpersonal competencies function similarly across different student populations or whether targeted interventions are needed for specific groups. In addition, future studies should extend the present findings by examining other factors that may influence psychological well-being, including personality traits, emotional intelligence, coping strategies, and social support. Given the potential conceptual overlap among the study variables, the use of factor analytic techniques, such as exploratory and confirmatory factor analysis, is also recommended to further establish the distinctiveness of assertiveness, social skills, and psychological well-being as separate constructs. Furthermore, longitudinal and experimental research designs are encouraged to better understand the directionality and developmental nature of these relationships, as the cross-sectional design of the current study limits causal interpretation.

8. Limitations of the Study

Several limitations should be acknowledged in interpreting the findings of this study. First, the use of a cross-sectional design limits the ability to establish causal relationships among the variables. The findings should therefore be interpreted as indicating statistical associations rather than causal effects. Second, the study relied on self-report questionnaires, which may be subject to response bias, including social desirability and self-perception bias. This may affect the accuracy of the responses provided by participants. Third, the sample was limited to first-year undergraduate students from selected public universities in Malaysia. Although the sample represents multiple geographical zones, the findings may not be fully generalizable to all university students, including those in private institutions or different academic levels. Finally, the study focused on a limited set of variables, namely assertiveness and social skills, in predicting psychological well-being. Other potentially important factors were not included in the model, which may explain the modest variance observed.

References

- Abdullah, K. (2005). *Tahap ketegasan pelajar universiti: Satu kajian di Universiti Teknologi Malaysia*. (Doctoral dissertation, Universiti Teknologi Malaysia).
- Alberti, R. E., & Emmons, M. L. (2017). *Your perfect right: Assertiveness and equality in your life and relationships* (11th ed.). New Harbinger Publications.
- Amin, R., Hussain, Z. N., Mahmood, S., Sadaf, A., & Saleem, M. (2021). Mediating role of social self-efficacy in linking positive thinking to satisfaction with life among first year university students. *Foundation University Journal of Psychology*, 5(1), 108-116.
- Ang, C. S. (2016). Assessment of social competence: exploring the dimensions of young adult social competence. *Jurnal Psikologi Malaysia*, 30(1).
- Arhuis-Inca, W., & Ipanaqué-Zapata, M. (2023). La relación entre las habilidades sociales, bienestar psicológico y rendimiento académico en estudiantado universitario de Chimbote, Perú. *Revista Electrónica Educare*, 27(2), 20-33. <http://dx.doi.org/10.15359/ree.27-2.15848>.
- Bamoussa, T. A. (2023). The level of assertiveness, Self-esteem and Communication Skills among Nursing Students of Helwan University. *Helwan International Journal for Nursing Research and Practice*, 2(4), 44-55. <https://doi.org/10.21608/hijnrp.2023.219291.1076>
- Betinelli, E. B., Rosa, J. C., de Oliveira Braz, A. L., Kormann, E., Zwierewicz, M., & da Silva, E. P. (2025). Educational social skills: A study on emotional intelligence in higher education. *Revista de Gestão Social e Ambiental*, 19(1), 1-18. <https://doi.org/10.24857/rgsa.v19n1-086>.
- Carneiro Machado, F., Braga Matias Santos, L., & Moreira, J. M. (2020). Social skills of Nursing and Psychology students. *Ciencias Psicológicas*, 14(1). <https://doi.org/10.22235/cp.v14i1.2131>
- Chakraborty, S., Ray, A., & Mani, S. K. (2020). A comparative study regarding assertiveness, self-esteem and leadership potentiality between students of selected nursing and general college in West Bengal. *International Journal of Nursing & Midwifery Research* (E-ISSN: 2455-9318), 7(3), 4-11. <https://doi.org/10.24321/2455.9318.202019>
- Chandio, K., Akhtar, P., Lolai, M. B., Sahoowal, A., & Dean, R. (2024). Association of assertiveness with emotional intelligence and self-esteem among undergraduate students. *Insights—Journal of Health and Rehabilitation*, 2(2), 193-199.
- Chaudhry, S., Tandon, A., Shinde, S., & Bhattacharya, A. (2024). Student psychological well-being in higher education: The role of internal team environment, institutional, friends and family support and academic engagement. *Plos one*, 19(1), e0297508. <https://doi.org/10.1371/journal.pone.0297508>
- Chelvarayan, A., Razak, A. A. B. A., & Fern, Y. S. (2023). Mental Health and Covid-19: The Malaysian Students' Perspective. *People and Behavior Analysis*, 1(1), 1-10. <https://doi.org/10.31098/pba.XXXX>
- Creswell, J. W. 2015. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New Jersey: Pearson.

- Demir, M., Jaafar, J., Bilyk, N., & Mohd Ariff, M. R. (2012). Social skills, friendship and happiness: A cross-cultural investigation. *The Journal of social psychology*, 152(3), 379-385. <https://doi.org/10.1080/00224545.2011.591451>
- Dougall, I., Vasiljevic, M., Kutlaca, M., & Weick, M. (2023). Socioeconomic inequalities in mental health and wellbeing among UK students during the COVID-19 pandemic: Clarifying underlying mechanisms. *PloS one*, 18(11), e0292842. <https://doi.org/10.1371/journal.pone.0292842>
- Durar, E., Keleş, İ., Durmuş, M., Ateş, K. (2024). The Relationship Between Problematic Internet Use, Psychological Well-Being and Assertiveness in University Students, *Adnan Menderes Üniversitesi Sağlık Bilimleri Fakültesi Dergisi*, 8(1), 29-41. <https://doi.org/10.46237/amusbfd.1289248>
- Eroğlu, İ. H., & Öner, H. (2023). The Mediating Roles of Emotional Expression and Self-Esteem in the Relationship Between Social Problem Solving and Psychological Well-Being of Nursing Students. *Celal Bayar Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 10(3), 222-231. <https://doi.org/10.34087/cbusbed.1104956>
- Ganji, F., Khani, F., Karimi, Z., & Rabiei, L. (2022). Effect of assertiveness program on the drug use tendency, mental health, and quality of life in clinical students of Shahrekord University of Medical Sciences. *Journal of Education and Health Promotion*, 11, 48. https://doi.org/10.4103/jehp.jehp_107_21
- García-Álvarez, D., Hernández-Lalinde, J., & Cobo-Rendón, R. (2021). Emotional intelligence and academic self-efficacy in relation to the psychological well-being of university students during COVID-19 in Venezuela. *Frontiers in psychology*, 12, 759701. <https://doi.org/10.3389/fpsyg.2021.759701>
- Gul, R., Batool, S., Khan, S. I., & Jabeen, F. (2023). The effects of social skills on academic competencies among undergraduate students. *Russian Law Journal*, 11(3S), 308-315.
- Habib, H., & Ahmad, S. (2024). An Exploratory Study: Lack of Assertiveness and its Relation with Aggression and Appearance Anxiety. *Research Journal of Psychology*, 2(3), 119-132.
- Haladin, N. A. B., Ibrahim, N. A., & Rajab, A. (2013). Assertiveness among undergraduates in a Malaysian public university: Implications towards improved employability. *Sains Humanika*, 65(2). <https://doi.org/10.11113/jt.v65.2346>
- Hayat, U., Zahra, K., Palwasha, A. U. S., Khan, K., Ahmad, M., & Iqbal, A. (2023). Impact of Social Anxiety on Assertiveness: Role of Gender. *Bulletin of Business and Economics (BBE)*, 12(3), 815-818. <https://doi.org/10.61506/01.00132>
- Hernández-Xumet, J. E., García-Hernández, A. M., Fernández-González, J. P., & Marrero-González, C. M. (2024). Exploring levels of empathy and assertiveness in final year physiotherapy students during clinical placements. *Scientific Reports*, 14(1), 13349. <https://doi.org/10.1038/s41598-024-64148-8>.
- Ibrahim, S. (2023). Life stressors and their relationship to assertiveness for university students. *Egyptian Journal of Social Work*, 15(1), 159-174. <https://doi.org/10.21608/ejsw.2023.175375.1168>
- Ilnytska, L. A., & Onischenko, I. O. (2022). Psychological features of assertiveness of students-psychologists. *Habitus*, (33), 92–97.

- Indreswari, H., Probawati, D., & Rachmawati, I. (2024). Psychological well-being and student academic burnout. *Jurnal Kajian Bimbingan Dan Konseling*, 7(3), 13. <https://doi.org/10.17977/um001v7i32022p138-149>
- Iqbal, T., & Nishat, A. (2017). Psychological Wellbeing of Adolescents: A Comparative Study. *Research Journal of Social Science & Management*, 6 (9), 23-41.
- Khazanah Research Institute. (2018). *The School-To-Work Transition Of Young Malaysians*. In Khazanah Research Institute (Issue December). [http://www.krinstitute.org/assets/contentMS/img/template/editor/20181212_SWTS_Presentation_DR_LIM .pdf](http://www.krinstitute.org/assets/contentMS/img/template/editor/20181212_SWTS_Presentation_DR_LIM.pdf)
- Khazanah Research Institute. (2020). *Work in an Evolving Malaysia. The State of Households 2020 Part II*.
- Kipli, M., Rahman, S., & Surat, S. (2022). Kebimbangan Dan Kemahiran Sosial Dalam Kalangan Pelajar Di Sebuah Institusi Pendidikan Islam Di Sarawak. *Jurnal Dunia Pendidikan*, 4(2), 49-62.
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610.
- Korda, M., Shulhai, A., Shevchuk, O., Shulhai, O., & Shulhai, A. M. (2025). Psychological well-being and academic performance of Ukrainian medical students under the burden of war: a cross-sectional study. *Frontiers in Public Health*, 12, 1457026. <https://doi.org/10.3389/fpubh.2024.1457026>
- Lazarus, A. A. (1971). *Behavior Therapy and Beyond*. New York: McGraw-Hil.
- Lazarus, A. A. (1973). Assertive behavior: A review and theoretical formulation. *Psychological Bulletin* 79(5): 376-391.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.
- Liaqat, S., & Deebea, F. (2025). Examining Pro-Social Behaviour and its Interplay with Psychological Well-being among Undergraduates. *The Critical Review of Social Sciences Studies*, 3(3), 2267-2278.
- Mahdavi, P., Valibeygi, A., Moradi, M., & Sadeghi, S. (2023). Relationship between achievement motivation, mental health and academic success in university students. *Community Health Equity Research & Policy*, 43(3), 311-317. <https://doi.org/10.1177/0272684X2110259>.
- Matud, M. P., López-Curbelo, M., & Fortes, D. (2019). Gender and psychological well-being. *International journal of environmental research and public health*, 16(19), 3531. <https://doi.org/10.3390/ijerph16193531>
- Moeller, R. W., & Seehuus, M. (2019). Loneliness as a mediator for college students' social skills and experiences of depression and anxiety. *Journal of Adolescence*, 73, 1-13. <https://doi.org/10.1016/j.adolescence.2019.03.006>
- Mohamad, F., Dahlan, R., Ismail, I. Z., Kadir, H., Kamaruddin, K. N., Shibraumalisi, N. A., & Shamsuddin, N. H. (2024). Self-system and mental health status among Malaysian youth attending higher educational institutions: A nationwide cross-sectional study. *Malaysian Family Physician: the Official Journal of the Academy of Family Physicians of Malaysia*, 19, 12. <https://doi.org/10.51866/oa.341>

- Mohammed Ibrahim Morsi, F., & Ezz-Eldin Prince, H. (2021). Assertiveness and Self-esteem of Final Year Nursing Students: Effect of Psycho-Educational Nursing Intervention. *Egyptian Journal of Health Care*, 12(1), 1186-1200.
- Moo, K. H., & Da Wan, C. (2023). Graduate employability in Malaysia: Unpacking the concept, policy and practices. *IIUM Journal of Educational Studies*, 11(2), 3-25.
- Nair, R. (2021). Perceived Stress and Assertiveness among Young Adults. *International Journal of Indian Psychology*, 9(2). <https://doi.org/10.25215/0902.210>
- Neuman, W.L. (2009). Social research methods: Qualitative and quantitative approaches (7th ed.). Boston, MA: Pearson/Allyn & Bacon.
- Norsayyidatina Che Rozubi. (2017). *Kesan program kesejahteraan psikologi menggunakan kaedah bermain kepada kanak-kanak sekolah* (Doctoral dissertation, Universiti Malaya).
- Pumpuang, W., Vongsirimas, N., & Klainin-Yobas, P. (2021). Do gender differences affect the psychological well-being of high schoolers in Thailand? *Journal of Population and Social Studies*, 29, 207-222. <http://doi.org/10.25133/JPSSv292021.013>
- Parray, W. M., Kumar, S., David, B. E., & Khare, S. (2020). Assertiveness predicts self-esteem, academic achievement, and stress: A study of Kashmiri adolescents. *Humanities & Social Sciences Reviews*, 8(1), 707-715. <https://doi.org/10.11648/j.rd.20220302>.
- Plantade-Gipch, A., Bruno, J., Strub, L., Bouvard, M., & Martin-Krumm, C. (2023). Emotional regulation, attachment style, and assertiveness as determinants of well-being in emerging adults. *Frontiers in Education*, 8, 1058519. <https://doi.org/10.3389/feduc.2023.1058519>
- Rahayu, M. N. M. (2023). Does Assertiveness Correlate with Second-Year Student Adjustment in the College? *Educational, Health and Community Psychology*, 12, 515-538. <https://doi.org/10.12928/jehcp.v1i2.25328>.
- Rosli, S. N., & Perveen, A. (2021). The relationship between prosocial behavior and psychological well-being among undergraduate students. *International Journal of Academic Research in Business and Social Sciences*, 11(6), 276-289. <https://doi.org/10.6007/IJARBS/v11-i6/10119>
- Sabrina, P., Ahmad, A., Salida, J., Azimah, M., & Annamma, K. (2022). Assertiveness and psychological empowerment among nursing students: A Malaysian perspective. *International Journal of Medical Science and Health Research*, 6(6). <https://doi.org/10.51505/ijmshr.2022.6602>
- Şen, S., Özkan, H. A., & Avşar, E. (2022). The Relationship Between Nursing Students' Psychological Empowerment and Assertiveness. *Journal of Academic Research in Nursing*, 8(1). <https://doi.org/10.55646/jaren.2022.92259>.
- Sial, Z. A., Naz, F. L., & Rasheed, A. (2021). Relationship Between Students' Social Skills and Academic Achievement at University Level. *VFAST Transactions on Education and Social Sciences*, 9(3), 232-241. <https://doi.org/10.21015/vtess.v9i3.754>
- Radley, K. C., & Dart, E. H. (2022). *Social skills teaching for individuals with autism*. Springer International Publishing.

- Rathus, S. A. (1973). A 30-item schedule for assessing assertive behavior. *Behavior Therapy*, 4(3), 398–406. [https://doi.org/10.1016/S0005-7894\(73\)80120-0](https://doi.org/10.1016/S0005-7894(73)80120-0).
- Rizkia Mulyana, F., Suherman, A., Juliantine, T., Blegur, J., Rohyana, A., & Saeful Bakhri, R. (2024). Building mental health and social skills: The positive impact of jigsaw model in taekwondo course. *Retos: Nuevas Perspectivas de Educación Física, Deporte y Recreación*, (61). <https://doi.org/10.47197/retos.v61.108372>
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>
- Ryff, C. D. (2018). Well-being with soul: Science in pursuit of human potential. *Perspectives on Psychological Science*, 13(2), 242–248. <https://doi.org/10.1177/1745691617699836>
- Salina, M. N., & Rahimi, N. M. (2015). Hubungan antara sokongan sosial dan kesejahteraan psikologi dalam kalangan pelajar universiti. *Jurnal Psikologi Malaysia*, 29(1), 1–12.
- Salleh, N. A., & Mustaffa, C. S. (2016). Gender differences in the psychological well-being among flood victims in Malaysia. *European Proceedings of Social and Behavioural Sciences*.
- Shengyao, Y., Xuefen, L., Jenatabadi, H. S., Samsudin, N., Chunchun, K., & Ishak, Z. (2024). Emotional intelligence impact on academic achievement and psychological well-being among university students: the mediating role of positive psychological characteristics. *BMC psychology*, 12(1), 389. <https://doi.org/10.1186/s40359-024-01886-4>
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons.
- Sekaran, U. (2005). *Research Methods for Business: A Skill Building Approach*. New Jersey: John Wiley and Sons
- Segrin, C. (2019). Indirect effects of social skills on health through stress and loneliness. *Health communication*, 34(1), 118-124. <https://doi.org/10.1080/10410236.2017.1384434>
- Soares, A. B., Monteiro, M. C., Medeiros, H. C. P., Maia, F. D. A., & Barros, R. D. S. N. (2021). Academic adaptation to the university: relationships among motivation, expectations and social skills. *Psicologia Escolar e Educacional*, 25, e226072. <https://doi.org/10.1590/2175-35392021226072>
- Suleman, N., Admani, A., Rahima, R., Ali, S. S., & Sami, A. (2022). *How do skills influence the students' employability in a developing economy?*
- Tekin, E. G. (2021). An investigation of psychological well-being, emotional intelligence and social well-being levels of university students. *Turkish Psychological Counseling and Guidance Journal*, 11(63), 567-575. <https://doi.org/10.17066/tpdrd.1051414>
- Vanhove, A., Opdecam, E., & Haerens, L. (2024). Fostering social skills in the Flemish secondary accounting education: perceived challenges, opportunities, and future directions. *Accounting Education*, 33(4), 414-449. <https://doi.org/10.1080/09639284.2023.2208106>.
- Wiersma, W. & Jurs, S.G. (2009). *Research Methods in Education: An Introduction*. (9th ed). Boston: Pearson Connexion Sdn. Bhd.

World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum.

Xiang, Y., Shuai, C., Zhang, Y., & Li, Y. (2024). The mental health of college students and the level of academic achievement: Knowledge about psychological well-being, self-control, and college learning. *Československá psychologie*, 68(1), 80-95. <https://doi.org/10.51561/cspsych.68.1.80>

Yan Carlo, Q. Q. (2022). Social Skills and academic performance in Educational Institutions of High Andean Puno-Peru in 2020. *Journal of Positive School Psychology*, 6(8).

Yelpaze, İ., Deniz, M. E., & Satıcı, B. (2021). Association between social connectedness and well-being: a study of the mediating role of psychological vulnerability. *Turkish Psychological Counseling and Guidance Journal*, 11(62), 367-382. <https://doi.org/10.17066/tpdrd.1001712>