Journal of Institutional Research South East Asia

JIRSEA

Volume 13 Number 1
2015

ISSN 1675-6061

Editor:
Nirwan Idrus PhD$_{Monash}$**IQA**$_{London}$

All papers are refereed by two appropriately independent, qualified experts and evaluated according to:

- Significance in contributing new knowledge
- Appropriateness for the Journal
- Clarity of presentation
- Technical adequacy

JIRSEA is indexed with the Directory of Open Access Journals, SCOPUS, EBSCoHost (Education Research Index/Education Research Complete) and currently applying to INFORMIT (Australasian Scholarly Research). As provided in the Ulrich's Directory, all journals in the DOAJ are also indexed in the Ulrich's Directory

SEAAIR
SOUTH EAST ASIAN ASSOCIATION FOR INSTITUTIONAL RESEARCH

http://www.seaairweb.info

CONTENTS

EDITORIAL BOARD

EDITORIAL

Not many of you realize that this is JIRSEA's 21st edition!
In many societies the *coming of age* is when one reaches his/her 21st birthday.

This edition is therefore a milestone for JIRSEA and SEAAIR the association that gave birth to JIRSEA. It is a coming of age indicating JIRSEA has gone through its childhood and adolescence and is now stepping onto adulthood.

It is most appropriate to give a note of thanks and appreciation to JIRSEA's inaugural editors Dr Gan Che Ng and Dr Raj Sharma who set the standard for JIRSEA. Raj of course was the impetus behind the founding of SEAAIR back in 2000.

We must not forget the members of the Editorial Board, the authors who entrusted their research outputs to JIRSEA and the article reviewers. There won't be any journal without them. These kind people had given their valuable time voluntarily.

On to this *coming of age* edition, we have contributors from Thailand, Malaysia, Japan and Australia discussing students' familiarization with university life, their writing ability enhancement by way of improving their grammar, their evaluation of their teachers' characteristics, as well as aspects of institutional research and error analysis. I add a little reflection in the Opinion Page especially given the ephemeral nature of things these days. It is inevitable that even education and its management must change with time. Those of us who believe in this wish that more educators and educationalists will see this point too soon.

Let's make the next 21 editions of JIRSEA just as successful. This can only be done by article contributions from you. Remember that JIRSEA is published twice a year, once in May/June and the second in October/November.

I look forward to receiving your articles. Refer to (http://wwwseaairweb.info) for details.

Happy reading!

Nirwan Idrus

Editor

THE IMPACT OF PERCEIVED TEACHER CHARACTERISTICS ON STUDENT EVALUATION: *A META-ANALYTIC STUDY*

Setsuko Mori
Faculty of Law, Kinki University, Japan

Abstract

One concern of qualitative rating metrics is rater bias. A halo effect is a well-researched phenomenon where a rater's opinion regarding only one aspect of the person being rated determines the overall rating. In order to determine whether a halo effect occurs in the completion of student evaluation of teaching (SET) ratings and estimate the magnitude of its effect, a meta-analysis of studies investigating the relationship between SET and a number of teacher personal characteristics was conducted. The teacher personal characteristics examined in the current study include students' perceptions of teacher personality, charisma, attractiveness, and perceived distance between teacher and students. This meta-analysis included 17 studies with 34 unique effect sizes and a total of 19,506 participants. The results suggest that 26% of the variance in SET can be explained by the teacher personal characteristics included in the study. Of the five teacher personal traits examined, teacher charisma had the largest impact on SET ratings followed by personality and attractiveness.

Keywords: student evaluation of teaching, halo effect, effects of teacher personal characteristics

Introduction

Since formal evaluation analysis began to be conducted in the early 50s in the United States of America, there has been a plethora of research addressing issues of validity of student evaluations (see Spoorenm et al., 2013 for an extensive summary).Some researchers (e.g. Feldman, 1977; Marsh, 1982; Zhao and Gallant, 2012) have found evidence that supports the validity of SET while others still have concerns (Greenwald, 1997). Of a variety of factors that can affect the validity of SETs, most research has investigated possible effects of students' expected grades and instructors' grading practices on students' ratings of instruction (e.g. Beran and Violato, 2005; Greenwald and Gillmore, 1997; Griffin, 2004; McPherson et al., 2009; Olivares, 2001; Remedios and Lieberman, 2008). Many researchers (e.g. Beren and Violato, 2005; Griffin, 2004; Maurer, 2006; McPherson et al.,2009; Olivares, 2001; Remedios and Lieberman, 2008) found a significant relationship between student's expected grades and SET ratings: The higher the expected grade, the higher SET. Other researchers (Greenwald

and Gillmore, 1997; Griffin, 2004; Olivares, 2001) contend that instructors can get higher SET ratings by following a more lenient grading policy.

In Japan, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) modified the standards for establishment of universities and from 2008 began to require universities to implement some kind of faculty development. As a result, many universities started to administer student evaluation of teaching (SET) as an integral component of faculty development. According to the Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2013), 93% of public and private universities now administer SETs in one form or another. In addition, 70% of those universities disclose the results either to both teachers and students or just teachers, and over 50% claim that they adopt a systematic approach to the use of SET results in order to improve the quality of teaching (MEXT, 2013). Nagahara, Sugawara, Matsuoka, and Ikeda (2011) conducted a survey on use of SETs at 95 Japanese universities, and reported that 73.7% of them said they administer SETs twice a year, and 57.9% said they administer SETs in all classes.

Generally speaking, SETs serve three purposes: (a) improving the quality of teaching, (b) providing information for instructor appraisal, and (c) providing evidence for institutional accountability (Spoorenm et al., 2013). In Japan, considering the fact that SETs are introduced as an essential part of faculty development, one could assume that the main purpose should be for instructors to gain insights into their strengths and weaknesses so that they can improve the quality of their teaching. However, since there is very little information regarding how SETs are actually used, a possibility that SETs are used as a tool for instructor appraisal or evidence for institutional accountability cannot be excluded.

Use of SET in instructor appraisal can be justified in a managerial approach where students are viewed as stakeholders, and meeting their demands for satisfaction plays an important part in teaching (Jauhianinen et al., 2009; Larsen, 2005 cited in Spoorenm et al., 2013). From this perspective, students' satisfaction is considered equivalent to learning or knowledge. However, there is most likely to be a major discrepancy between what teachers and students regard as effective teaching. For this reason, many instructors may doubt the validity and reliability of SET results. After all, formative use of SET (e.g. as feedback to improve teaching) and summative use (e.g. as a tool to make administrative decisions) have potentially contradicting purposes. Lack of transparency regarding the use of SET data does not seem to help reduce Japanese instructors' skepticism.

Given the apparent ubiquity of SET in higher education in Japan, and presumably other parts of Asia, and the possibility that the results of SET will be used in high-stakes personnel decisions, it is important to investigate exactly what is being measured and the reliability of such measurements. In the field of second/foreign language education, there has been almost no research done dealing with the validity and reliability of SETs (Mori and Tanabe, 2012; and Tanabe and Mori, 2013, are exceptions). Since the findings of Mori and Tanabe (2012) and Tanabe and Mori (2013) imply that students in language classrooms react in similar ways as they do in classes of other disciplines, it is hoped that the present meta-analysis will serve as a starting point, blueprint, and call to action for more in-depth investigation into SET in the field of foreign language education in Asia.

Halo Effects on SET Ratings

The halo effect, first coined by Edward Thorndike (1920), refers to a phenomenon where a rater's opinion regarding only one aspect of the teacher determines the rest of the rating. According to Nisbett and Wilson (1977), the halo effect may be a result of extending an overall impression onto unknown or ambiguous characteristics (e.g. if we like certain qualities of people, we suppose we would also like unknown qualities), and/or may be so powerful that we change our perceptions of known characteristics (e.g. we may rate an attractive person as being more intelligent). In addition to those two possible explanations of halo effects, Feeley (2002) argues that the halo effect may occur due to students' lack of effort.

Sources of halo effects can include students' perception of a teacher's personal traits such as personality, charisma, and physical attractiveness. Although Felder (1995) denies the long-lasting skepticism among educators that SETs are a popularity contest, many other researchers found a strong correlation between students' perceived teacher personality and their evaluation of instruction. Feldman (1986), for instance, reviewed and synthesized the extant research correlating university teachers' personality characteristics with the teachers' effectiveness, and concluded that a wide variety of personal traits are strongly associated with rated effectiveness of teaching. His contention is supported by more recent studies.

Clayson and Sheffet (2006) conducted survey research using the Five Factor Model of Personality often referred to as the Big Five (Digman, 1990). The Big Five represent five dimensions of personality, namely agreeableness, conscientiousness, stability, extroversion, and creativity (openness). The study found a consistent and positive relationship between those personality traits and course evaluations, and the association was formed within fewer than five minutes of initial contact and grew stronger over the term. Clayson (2013) later conducted a similar study and confirmed the previous finding that showed students' first impressions of the instructor influenced the final evaluation given in a class. This finding is in congruence with some evidence presented by Ortinau and Bush (1987), and Sauber and Ludlow (1988) indicating that subsequent class experience may do little to change students' initial impressions of the instructor and teaching effectiveness.

Patrick (2011) also examined whether personality traits measured by the perception of students could predict student evaluations of teachers and courses using a similar Big Five Inventory (John et al., 1991 cited in Patrick 2011). He found agreeableness, conscientiousness, extroversion, and openness correlated positively and neuroticism correlated negatively with student evaluations of the teachers and the courses.

Experimental studies where classroom conditions were manipulated have found remarkable effects of perceived teacher personality on the evaluations. Widmeyer and Loy (1988), for example, examined "warm/cold" manipulation on first impressions of teacher and teaching ability. In this experimental study, half of the participants were told that the stimulus person had a "warm" personality whereas the other half were told that he had a "cold" personality before he appeared. It was found that the former group perceived him as a more effective teacher than the latter group.

Other researchers investigated possible influences of different teacher personal traits on SETs. Shelvin, Banyard, Davies, and Griffiths (2000) looked at how influential instructors' charisma was on teaching evaluations and found that the charisma factor explained approximately 70%

of the variance in the lecture ability. Delucchi (2000),and Gurung and Vespia (2007) both included a likability factor in their analysis, and found that teachers who were rated high in likability earned high ratings. Instead of likability, Feeley (2002) explored a degree of closeness or distance between two people referred to as nonverbal immediacy. The results indicated that nonverbal immediacy was strongly associated with teaching evaluations. Wendorf and Alexander (2005) examined the influence of students' perception of fairness on their satisfaction with the instructors. Their study revealed that instructor satisfaction was significantly affected by the perception of the fairness of instructor-student interactions.

There have also been many studies that investigated the influence of perceived physical attractiveness on students' evaluations of teachers (e.g. Ambady and Rosenthal, 1993; Goebel and Cashen, 1979; Hammermesh and Parker, 2005). Research interest in this area goes as far back as to the 70s. Goebel and Cashen (1979) showed black-and-white photographs of teachers to students in grades 2, 5, 7, 11, and 13, and had them rate teacher performance on seven factors. The results showed that across all the developmental levels, ratings of unattractive teachers tended to be lower on all factors. Hammermesh and Parker (2005) also found that the impact of beauty on instructional ratings was significant.

Recently, online professor rating systems have enabled researchers to utilize much larger databases in exploring the role of physical attractiveness on SET. Especially, RateMyProfessor.com, the largest online professor rating site, has been used as the major source in many analyses (e.g. Felton et al., 2004; Riniolo, et al., 2006; Felton, et al., 2008; Freng and Webler, 2009). According to RateMyProfessors.com, 14 million ratings, 1.3 million professors and 7,000 schools in the United State, Canada and the United Kingdom, have been added to their site. More than four million college students use RateMyProfessors.com on a monthly basis. On RateMyProfessor.com, together with such categories as clarity, helpfulness, and easiness, students rate their professors on hotness on a binary score. In this category, professors are awarded chili peppers based on the sum of positive and negative (hot or not) ratings. Using this hotness scale, Riniolo, Johnson, Sherman, and Misso (2006), Felton, Koper, Mitchell and Stinson, (2008), and Freng and Webler (2009) all found that professors who were perceived as attractive received higher student evaluations.

Purpose of the Study

Despite the fact that most English classes and teachers are also subjects of SETs, to date there has been almost no research done dealing with the validity of SETs in the field of second/foreign language education (Mori and Tanabe, 2012; and Tanabe and Mori, 2013 are exceptions). Therefore, the author designed the current study to provide a comprehensive overview of previous studies related to halo effects on SETs. Meta-analysis of studies on the topics of interest will enable educators and researchers as well as policy-makers to examine a wealth of research and understand possible biases of SETs. More specifically, meta-analyses were examined to determine the overall relationship between student perception of teacher personal characteristics and SET ratings, and identify which moderating variables (i.e., personality, charisma, attractiveness, and distance from students) account the most for variations in SET ratings.

Method

Identification of Studies

Using online databases, namely Academic Search Premier and ERIC, an extensive literature review was conducted to identify published studies that addressed some aspects of halo effects on SETs. Search terms used included: SET, student evaluation, student evaluation of teaching, halo effect, teacher personality, teacher effectiveness, and attractiveness. Articles specific to relationships between teacher personal characteristics such as personality, charisma, attractiveness and distance from students and SET ratings were selected from the large sample of studies.

Criteria for Inclusion of Studies

The criteria for being included in the meta-analyses were as follows: (1) studies were peer-reviewed, (2) studies must have provided appropriate statistical data such as means, standard deviations and correlations, (3) studies included college students' evaluation of teaching, (4) studies used overall scores of instruction or teaching effectiveness as a dependent variable, and (5) studies investigated students' perceived teacher personal characteristics on their evaluation of teaching.

Identification of Possible Moderating Variables

Table 1 indicates that five teacher personal characteristics, namely personality, charisma, attractiveness, distance from students, and others (i.e. attitude, image and fairness)were identified as moderating variables. Considering multidimensionality of personality, six personality types were included in subsequent analysis. There has been a general consensus that an "adequate taxonomy for personality attributes" could be created by five factors (Digman, 1990). These types are referred to as the Big Five. However, ideas of what constitutes the big five personality attributes differ slightly depending on researchers. Clayson and Sheffet (2006), for example, included agreeableness (i.e. friendly, trusting, cooperative), conscientiousness (i.e. methodical, well-organized, respectful), emotional stability (i.e. relaxed, less emotional, less prone to distress), extroversion (i.e. seeking out the company of others, energized by interactions), and creativity/openness (i.e. open-minded, creative, interested in culture). On the other hand, Patrick (2011) included neuroticism (i.e.instable, anxious, moody) instead of stability. Although emotional stability and neuroticism can be considered asimilar dimension, they were analyzed separately in the current study.

A charisma factor was examined in Harvey, et al. (2003), Huang and Lin (2014), and Shelvin et al. (2000) which showed that it has an impact on people's judgment of others, and is suspected to be a salient trait in students' perceptions of teachers in assessing their teaching effectiveness. Five studies (Feeley, 2002; Felton, et al., 2008;Freng and Webber, 2009;Gurung and Vespia, 2007;Riniolo, et al., 2006) explored an attractiveness factor in relation with SET. Except for Gurung and Vespia (2007),which used their original measurement for attractiveness, four studies obtained data from RateMyProfessor.com. On RateMyProfessor.com, attractiveness is measured as the sum of positive and negative (hot or not) ratings. The concept of distance from students was operationalized as amicability (Sherman et al., 1975), likability (Delucchi, 2000; Gurung and Vespia, 2007), nonverbal immediacy (i.e., a degree of closeness or distance between people defined in Feeley, 2002), self-disclosure (measured in terms of intent, amount, positiveness, depth and honesty in Lannutti and Strauman, 2006) and approachability (Gurung and Vespia, 2007).

Concepts of teacher attitude, student image of the teacher, and student idea of fairness were also included in the current analysis, as they have shown to have an impact on SETs. However, they were labeled as others because only one study was available for analysis on each topic. Specifically, attitude was measured by such items as "good attitude toward class," "showed respect toward students," and "pleasant to deal with" in Kim et al (2000). Image compatibility (Dunegan and Hrivnak, 2003) refers to closeness or distance between student's mental images of an "ideal" instructor and the current instructor in terms of characteristics, skills and qualifications. Although Wendorf and Alexander (2005) included three types of fairness, distributive, procedural and interactional, only interactional fairness was included in analysis of the current study as the first two refer to fairness of an outcome and rules whereas interactional fairness refers to the quality of interpersonal treatment, which is pertinent to the study.

Results

Based on the criteria for inclusion in this study, 17studies were identified (see Table 1). The following variables were examined among the 17 studies: student perception of teacher personality (n=4), charisma (n=3), attractiveness (n=5), distance from students (n=5) and others (n=3).The total sample size was 19,506 participants.

Meta-analysis is a statistical method used to combine results of individual studies, and can be best understood by their effect sizes. Effect size is a measure of strength of a phenomenon. To interpret effect sizes, Cohen's (1988) guidelines for correlation effects and mean difference effects are most commonly used. According to Cohen's guidelines for correlation effect sizes, r=.10 is a small effect, r=.30 is a medium effect, and r=.50 or above indicates a large effect. This means a small effect (r=.10), a medium effect (r=.30) and a large effect (r=.50) indicate that 1%, 9% and 25% of the variance in one variable is accounted for the other variable. Cohen's guidelines for the standardized mean difference between two groups are as follows: d=.20 is a small effect, d=.50 is a medium effect, and d=.80 is a large effect. This study followed Cohen's guidelines in interpreting the effect sizes. In a meta-analysis with 17 studies, all except Patrick (2011), Delucchi (2000)and Lannutti and Strauman (2006) found a medium-to-large effect size between some aspects of teacher personal characteristics and SET (Table 1).

A meta-analysis was conducted to combine SET feedback with student perceived teacher personal characteristics. The homogeneity test was significant (Q=6658.240, df=31, p=0.000), which means that the null hypothesis that all studies share a common effect size was rejected. To address the variation across the studies, a random-effects meta-analysis, rather than a fixed-effects meta-analysis, was performed. Using random-effects weights, the summary estimate of the correlation is 0.511 with a 95% confidence interval of 0.426 to 0.588. The Z-value is 10.060, and the p-value is <0.000.This means that 26% of the variance in the SET ratings can be explained by student-perceived teacher personal characteristics.

Table 1: Summary of Studies Included in Meta-analysis

Characte-ristics	Author(s)	n size	# of Dimen-sions	Dimensions	*R*	Summary
Persona lity	Clayson & Sheffet (2006)	481	5	conscientiousness creativity (openness) agreeableness stability extroversion	0.65	Students' perceived instructor personality showed (Big Five) significant correlations with SET.

Characte-ristics	Author(s)	n size	# of Dimen-sions	Dimensions	*R*	Summary
	Murray (1975)	35	1	extroversion	0.51	Extroversion was significantly correlated with overall teacher rating.
	Patrick (2011)	176	5	agreeableness openness conscientiousness extroversion neuroticism	0.22	Students' perceived instructor personality (Big Five) showed significant correlations with SET.
	Radmacher & Martin (2001)	351	1	extroversion	0.79	Teachers' extroversion was correlated highly with students' evaluations of teaching effectiveness.
Charisma	Harvey, Royal & Stout (2003)	117	1	charisma	0.80	Charisma factor was highly correlated with students' perception of instructor performance.
	Huang & Lin (2014)	1078	1	charisma	0.60	There was a strong association between teacher charisma and student evaluation of teaching.
	Shevlin, Banyard, Davies & Griffiths (2000)	199	1	charisma	0.83	Charisma factor accounted for appropriately 70% of the variation in the lecture ability.
Attractiveness	Feeley (2002)	128	2	attractiveness	0.50	Attractiveness was strongly associated with teaching evaluations.
	Felton, Koper, Mitchell & Stinson (2008)	6852	1	attractiveness	0.64	Professors perceived as attractive received higher student evaluations.
	Freng & Webber (2009)	2281	1	attractiveness	0.37	Professors perceived as attractive received higher student evaluations.
	Gurung & Vespia (2007)	861	1	attractiveness	0.50	Teachers who were perceived as good-looking received higher SET.
	Riniolo, Johnson, Sherman & Misso (2006)	1714	1	attractiveness	0.53	There was a significant relationship between student's perception of instructor attractiveness and SET.
Distance from students	Delucchi (2000)	205	1	likability	0.22	Students who rated their instructor high in likability rewarded that instructor with high ratings in overall teaching ability.
	Feeley (2002)	128	1	non verbal immediacy (closeness between two people)	0.50	Nonverbal immediacy was strongly associated with teaching evaluations.
	Gurung & Vespia (2007)	861	2	approachability likability	0.50	Teachers who were perceived as likable and approachable received higher SET.
	Lannutti & Strauman (2006)	333	5	self-disclosure intent self-disclosure amount self-disclosure positiveness self-disclosure depth self-disclosure honesty	0.21	Perceptions of the intentionality, positiveness, and honesty of instructor self-disclosure were significantly and positively correlated with evaluations of the instructor.
	Sherman & Blackburn (1975)	1500	1	amicability	0.58	Professors perceived as amicable received higher teaching competence ratings.
Others	Kim, Damewood & Hodge (2000)	1504	1	attitude	0.48	Students who perceived positive attitudes exhibited by professors evaluated them higher in overall teaching effectiveness.
	Dunegan & Hrivnak (2003)	127	1	image compatibility	0.30	SET scores were significantly correlated with image compatibility (i.e., the difference/closeness between an image of an "ideal" instructor and an image of the instructor in this course).
	Wendorf & Alexander (2005)	575	1	interactional fairness (a quality of interpersonal treatment)	0.74	Instructor satisfaction was significantly affected by the perception of the fairness of instructor-student interactions

Subsequently, to more clearly understand the effects that each moderating variable had on SETs, the 17 studies were combined and analyzed according to different dimensions of teacher personal characteristics (Table 2).

Table 2:Overall Meta-Analysis Results by Teacher Personal Characteristics

Meta-analysis topic compared with SET	Number of dimensions combined	Pearson's *r* effect size	Percent of variance explained	Size of effect
Personality	12	0.50	25%	Large
Charisma	3	0.76	58%	Large
Attractiveness	5	0.56	31%	Large
Distance	4	0.25	6%	Small-Medium
Others	3	0.54	29%	Large
Overall total	27	0.51	26%	Large

Note: Size of effects are based on Cohen's guidelines for r, which are: .1=small, .3=medium and .5 large.

Comparing the effect sizes according to teacher personal characteristics, the range for Pearson *r* values was between 0.24-0.76, suggesting all but one teacher personal variables had a large effect size. Alternatively stated, 6-58% of the variance in the variables of interest is accounted for by student evaluation of teaching. Of the five teacher personal variables, charisma had the largest effect size followed by attractiveness, others and personality.

Table 3:Results from Combined Studies by Teacher Personality

Dimensions of personality	Number of dimensions combined	Pearson's *r* effect size	Percent of variance explained	Size of effect
Agreeableness	2	0.63	40%	Large
Conscientiousness	2	0.35	12%	Medium-Large
Extroversion	4	0.53	28%	Large
Neuroticism	1	-0.31	10%	Medium
Openness	2	0.64	41%	Large
Stability	1	0.64	41%	Large

In order to identify which types of teacher personality have a greater impact on student evaluations of teaching, the effect sizes according to different personality dimensions were compared (Table 3).All personality types had medium or larger effect sizes. Conscientiousness

and neuroticism had medium or medium to large effect sizes while agreeableness, extroversion, openness, and stability had large effect sizes.

Discussion and Conclusion

The main purpose of this study was to determine whether a halo effect occurs in the completion of SET ratings and to estimate the magnitude of its effect. Specifically, this study focused on the possible effects of students' perception of teacher characteristics on SET ratings. The results of the meta-analysis clearly indicate that students' perception of teacher personal traits has a strong impact on their evaluation of instruction. As a matter of fact, the effect accounts for approximately 25% of the variation of student's perception of teaching effectiveness.

In order to identify what aspects of teachers have more effects on SETs, the 17 studies were meta-analyzed with teacher personal characteristics as moderating variables. The results show that students' perceived teacher charisma, personality and attractiveness had large effect sizes. Of these three variables, charisma had the largest effect size. Although charisma was dealt with as an independent construct in the studies included in this analysis, a possible explanation for this finding is that charisma may be embodied by certain personality traits and attractiveness. Additionally, the studies that explored an aspect of attractiveness relied on dichotomous ratings to measure instructors' hotness (hot or not hot). Although the objective of this study is not to present a definitive definition ofa charismatic or hot teacher, it can be suspected that charisma or hotness (attractiveness) cannot be separated from certain personality traits.

As for personality traits, the results of the analysis indicated that students tended to rate agreeable, extroverted and open teachers higher on overall teaching. Although it is not the purpose of this study to convince teachers to be more agreeable, extroverted and open so that they can receive higher overall ratings, there is a possibility that teachers with these positive personality traits may be effective teachers (i.e. easy to ask questions of, open to student questions and opposing ideas). Nevertheless, if the impression of the instructor that the students form five minutes into the first class determines their final evaluation of instruction, as suggested by Clayson (2013), some teachers considered introverted or less open may never receive high evaluations regardless of what they do or not do in class. If that is the case, it raises questions regarding the validity and interpretation of student evaluation of teaching. Although the strong relationship between non-instructional variables such as personality and attractiveness and SET ratings does not in itself invalidate the instruments as far as students' perception reflects important components of instruction, the problem occurs when the evaluations are used blindly as a single measure of teaching quality. Therefore, student evaluations of teaching can be best used together with multiple sources of data, and should be closely monitored both by faculty and administrators when they are used as indicators of effective teaching. Furthermore, as Wright and Guarnieri (2012) suggest, it seems beneficial to pair SET results with some forms of consultation including extended discussion of teaching effectiveness, and training or workshops, instead of dumping all the SET results on individual teachers.

The results of the current study further our understanding of possible biases on students' evaluation of instructions, but it has limitations. One obvious limitation is the number of studies

meta-analyzed. Although there is no clear definition of how many studies should be considered suitable for a meta-analysis, increasing the number of studies on this topic will certainly increase generalizability of the findings. Another limitation inherent in any meta-analysis research is that "meta-analysis adds together apples and oranges" (DeCoster, 2004). In the case of this study, various types of scales were used to measure student evaluations of teaching, and teachers' personal traits. Consequently, a possibility that student evaluations varied based on the way they were measured cannot be completely denied.

For future research, it would be interesting to investigate the relationship between SET results and learning. Wright and Guarnieri (2012) found that there is a medium effect size between SET ratings and student achievement as a result of analysis of six meta-analysis studies. However, student achievement in those studies was measured by final exams or final course grades. Consequently, whether the results were derived from reported biases where a lenient grading policy may result in high SET ratings or actual learning cannot be determined. As a topic for future studies, it would be interesting to investigate the correlation between SET ratings and student willingness to work, or work completion. This begs the question:do students work more or harder for more popular teachers? In short, more research examining the relationship between SET ratings and actual learning will deepen our understanding of possible biases and enhance debates regarding the validity of student evaluations of effective teaching.

Furthermore, Mori and Tanabe (2012) indicated that there was no significant difference between English classes and other subject classes in terms of the effects of non-instructional variables on SETs, more investigation on this line of research will be necessary to verify that a strong association between teacher characteristics and SET ratings is not subject dependent. One of the limitations not mentioned in the previous paragraph is the lack of generalizability of this meta-analysis. All of the studies in the meta-analyses are concerned with American college students. This is unfortunate but inevitable, given the lack of studies in other settings. It is hoped that this paper will stimulate and act as a guideline for research in Asian settings. This is particularly important, given the ubiquity of SET at the university level, the paucity of relevant studies, and the importance that SETs are acquiring at many universities.

References

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431-441.

Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment and Evaluation in Higher Education, 30*, 593-601.

Clayson, D. E. (2013). Initial impressions and student evaluation of teaching. *Journal of Education for Business, 88*(1), 26-35.

Clayson, D. E., & Sheffet, M. J. (2006). Personality and student evaluation of teaching. *Journal of Marketing Education, 28*, 149-160.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. 2nd ed*. New York: Academic Press.

DeCoster, J. (2004). Meta-analysis notes. Retrieved June 8, 2014 from http://www.stat-help.com/notes.html.

Delucchi, M. (2000). Don't worry, be happy: Instructor likability, student perceptions of learning, and teacher ratings in upper-level sociology courses. *Teaching Sociology, 28*, 220-231.

Digman, J. M. (1990). Personality structure: An emergence of the five factor model. *The Annual Review of Psychology, 41*, 417-440.

Dunegan, K. J., & Hrivnak, M. W. (2003). Characteristics of mindless teaching evaluations and the moderating effects of image compatibility. *Journal of Management Education, 27*, 280-303.

Feeley, T. H. F. (2002). Evidence of halo effects in student evaluations of communication instruction. *Communication Education, 51*, 225-236.

Felder, R. M. (1995). What do they know anyway? *Chemical Engineering Education, 26*(3), 134-135.

Feldman, K. A. (1977). Consistency and validity among college students in rating their teachers and courses: A review and analysis. *Research in Higher Education, 6*, 223-274.

Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis. *Research in Higher Education, 24*(2), 139-213.

Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: Relations between perceived quality, easiness, and sexiness. *Assessment and Evaluation in Higher Education, 29*(1), 91-108.

Felton, J., Koper, P., Mitchell, J., & Stinson, M. (2008). Attractiveness, easiness and other issues: Student evaluations of professors on Ratmyprofessors.com. *Assessment & Evaluation in Higher Education, 33*(1), 45-61.

Freng, S., & Webber, D. (2009). Turning up the heat on online teaching evaluations: Does "hotness" matter? *Teaching of Psychology, 36*, 189-193.

Goebel, B. L., & Cashen, V. M. (1979). Age, sex, and attractiveness as factors in student ratings of teachers. *Journal of Educational Psychology, 71*(5), 646-653.

Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. American Psychology, 52, 1182-1186.

Greenwald, A. G., & Gillmore, G. M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology, 89,* 743-751.

Griffin, B. W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology*, *29*, 410-425.

Gurung, R., & Vespia, K. (2007). Looking good, teaching well? Linking liking, looks, and learning. *Teaching of Psychology*, *34*, 5-10.

Hamermesch, D. S., & Parker, A. (2005). Beauty in the classroom: Instructor's pulchritude and putative pegagogical productivity. *Economics of Education Review, 24*, 369-376.

Harvey, S., Royal, M., & Stout, D. (2003). Instructor's transformational leadership: University student attitudes and ratings. *Psychological Reports, 92*, 395-402.

Huang, Y. C., & Lin, S. H. (2014). Assessment of charisma as a factor in effective teaching. *Educational Technology & Society, 17*(2), 284-295.

Jauhianinen, A., Jauhiainen, A., & Laiho, A. (2009). The dilemmas of the "efficiency university" policy and the everyday life of university teachers. *Teaching in Higher Education, 14*, 417-428.

Kim, C., Damewood, E., & Hodge, N. (2000). Professor attitude: Its effect on teaching evaluations. *Journal of Management Education, 24*, 458-473.

Lannutti, P. J., & Strauman, E. C. (2006). Classroom communication: The influence of instructor self-disclosure on student evaluations. *Communication Quarterly, 54*(1), 89-99.

Marsh, H. (1982). Validity of students' evaluations of college teaching: A multitrait-multimethod analysis. *Journal of Educational Psychology, 74*, 264-279.

Maurer, T. W. (2006). Cognitive resonance or revenge? Student grades and course evaluations. *Teaching of Psychology, 33*, 176-179.

McPherson, M. A., Todd, J. R., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal, 35*, 37-51.

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2013) Daigakuniokeru kyouikunaiyoutouno kaikakujoukyoutou ni tsuite.Available athttp://www.mext.go.jp/a_menu/koutou/daigaku/04052801/__icsFiles/afieldfile/2014/03/10/1341433_03.pdf

Mori, S., &Tanabe, Y. (2012).Effects of perceived teacher personality on student class evaluation.*Kinki University Center for Liberal Arts and Foreign Language Education Journal (Foreign Language Edition)*,*3*(1), 163-181.

Murray, H. G. (1975). Predicting student ratings of college teaching from peer ratings of personality types. *Teaching of Psychology, 2*(2), 66-70.

Nagahara, K., Sugawara, R., Matsuoka, K., &Ikeda, H. (2011). A questionnaire survey of course evaluation conducted in Japanese universities. *Bulletin of Hokkaido Bunkyo University*,*12*,157-172.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgment. *Journal of Personality and Social Psychology. 35*, 250-256.

Olivares, O. J. (2001). Student interest, grading leniency, and teacher ratings: A conceptual analysis. *Contemporary Educational Psychology, 26*, 382-399.

Orinau, D. J., & Bush, R. P. (1987). The propensity of college students to modify course evaluations and its impact on course performance information. *Journal of Marketing Education, 9*, 42-52.

Patrick, C. L. (2011). Student evaluations of teaching: Effects of Big Five personality traits, grades, and validity hypothesis. *Assessment & Evaluation in Higher Education, 36*, 239-249.

Racmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology, 135*(3), 259-268.

Remedios, R., & Lieberman, D. A. (2008). I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal, 34*, 91-115.

Riniolo, T. C., Johnson, K. C., Sherman, T. R., & Misso, J. A. (2006). Hot or not: Do professors perceived as physically attractive receive higher student evaluations? *Journal of General Psychology, 133*, 19-35.

Sauber, M. H., & Ludlow, R. R. (1988). Student evaluation stability in marketing: The importance of early class meetings. *The Journal of Midwest Marketing, 3*(1), 41-49.

Sherman, B. R., & Blackburn, R. T. (1975). Personal characteristics and teaching of college faculty. *Journal of Educational Psychology 67* (1): 124–31.

Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation in higher education: Love me, love my lectures? *Assessment & Evaluation in Higher Education, 25*, 397-405.

Spooren, P., Brockx, B., & Mortelman, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research, 83*(4), 598-642.

Tanabe, Y., &Mori, S. (2013). Effects of perceived teacher personality on student class evaluations: A comparison between Japanese instructors and native English speaking instructors. *International Journal of English Linguistics,3*(3), 53-65.

Wendorf, C. A., & Alexander, S. (2005). The influence of individual- and class-level fairness-related perceptions on student satisfaction. *Contemporary Educational Psychology, 30*, 190-206.

Widmeyer, W. N., & Loy, M. L. (1988). When you're hot, you're hot! Warm-cold effects in first impressions of persons and teaching effectiveness. *Journal of Educational Psychology*, *80*, 118-121.

Wright, S. L., & Jenkins-Guarnieri, M. A. (2012). Student evaluations of teaching: Combining the meta-analyses and demonstrating further evidence for effective use. *Assessment & Evaluation in Higher Education, 37*(6), 683-699.

Zhao, J., & Gallant, D. J. (2012). Student evaluation of instruction in higher education: Exploring issues of validity and reliability. *Assessment & Evaluation in Higher Education, 37*, 227-235.

MOTIVATIONAL ORIENTATION, PERCEIVED STRESS AND UNIVERSITY ADJUSTMENT AMONG FIRST YEAR UNDERGRADUATES IN MALAYSIA

P. C. Siah
Department of Psychology and Counseling, Universiti Tunku Abdul Rahman, Malaysia

and

S. H. Tan
Faculty of Education, Universiti Teknologi Malaysia, Malaysia

Abstract

The aims of this study are to investigate the relationship between motivational orientations, perceived stress and university adjustment among the first year undergraduates in Malaysia. A total of 200 questionnaires were distributed to the respondents by using a purposive sampling method. The results found significant positive relationships between levels of self-determination with university adjustment but negative relationships with perceived stress, but the positively relationships between extrinsic motivation identified regulation and both the college adaptation and perceived stress were unexpected. Based on the results, we suggested programs could be designed to make the public aware about the importance of intrinsic motivation in the choice of higher education. Parents should allow their children to select the subjects their own rather than what the parents thought is good for them.

Keywords: first year undergraduates; self-determination; university adjustment; perceived stress

Introduction

The Malaysian higher education sector has undergone substantial growth as a result of efforts taken by the Ministry of Education to expand the education industry. It is the government's long-term goal to make Malaysia a regional center of excellence in education. The growth of higher education in Malaysia can be seen in increasing in students' enrolment (Ariffin, Ahmad, Ahmad, & Ibrahim, 2008). The statistics on tertiary education in Malaysia showed an increasing trend

over a ten-year period from 2001 to 2010. The number of students enrol has been increased from 628,479 in 2001 to 1134,134 in 2010 (Ministry of Higher Education, 2012) . This means that more and more people in Malaysia are taking the decision to further their academic qualification beyond the secondary education.

Attending university for the first time is a crucial test for is a test of the individual's capacity to adjust to a new and foreign environment (Dyson & Renk, 2006). The freshman year has been described as a critical year as it provides the foundation for the subsequent years of study and persistence (Reason, Terenzini, & Domingo, 2006). Being new to the environment in the university, most first year undergraduates may be unsure about whether they can meet the expectations they plan for themselves and also those imposed on them by their families and friends. Thus, a majority of the first year undergraduates go through an adjustment phase upon entry to university, with each student varying in his or her pace of development (Dyson & Renk, 2006).

According to Colman (2001), stress is referred to psychological and physical strain or tension caused by physical, emotional, social, economic, or occupational circumstances, event, or experiences in which an individual faces difficulty to manage or endure. Stressor is defined as events or social environments that place pressure on an individual and cause him or her to experience stress. Not all stress affects individuals negatively. When it is under right condition, stress brings positive effects, in which it enables individuals to respond effectively in a challenging situation (Elias, Ping, & Abdullah, 2011). Additionally, when under same situation and condition, the degree of stress experience by an individual may be different from another individual (Robotham & Julian, 2006).

Most first year students reported that their perceived stress in university life is higher than they anticipated (Compas, Slavin, Wagner, & Vannatta, 1986), as they not only face the pressure relevant to home sick and adjust to a new social environment, but also the challenges associated with home-sickness and the adjustment to a new social environment (Ross, Niebling, & Heckert, 1999). A study conducted in Malaysia University found that the most important causes of stress reported by the first year undergraduates are financial issues due to financial crisis that affected the entire world and the rise in cost living, lack of sleep and family problem (Redhwan, Sami, Karim, Chan, & Zaleha, 2009). Other less important causes of stress reported were loneliness, conflict with own belief, poor time management, unhealthy diet, lack of exercise, noisy pollution, stressful life events such as accidents and death of relatives and friends. It is estimated that about 25% will drop out in the first year (Arnett, 2004), and most withdrawn from their study in the first two years are due to the lack of adjustment to the new environment (Kerr, Johnson, Gans, & Krumrine, 2004).

University adjustment can be divided into academic adjustment, social adjustment, and emotional adjustment (Baker & Siryk, 1984). Academic adjustment reflects the student's opinion about the quality of the education that they are receiving and the degree of satisfaction with their academic performance. Social adjustment refers to the student's reactions to the social opportunities on campus and their success in interacting and communicating with people. Lastly, emotional adjustment reflects the student's experiences of homesickness. The student's overall success and satisfaction in their university life is determined by a successful adjustment made in these three areas (Conti, 2000). Therefore, adjustment is a multi-dimensional process of

interaction between an individual and his or her environment. College or university adjustment is referred to as multi-faceted, which involves an order of demands varying in kind and degree and which needs a diversity of responses and adjustment strategies (Baker & Siryk, 1984).

Many factors are found to be relevant to the university adjustment, such as academic aptitude, social integrating, family income, gender and parental support (Brooks & DuBois, 1995). Studies showed that mental health variables, such as stress and depression, are more important factors than GPA to predict university adjustment (Wintre & Yaffe, 2000). It has shown that nearly 80% of college students reported the experience of stress (Rawson, Bloomer, & Kendall, 1994). Persistent high levels of stress may affect the mental and physical health of some first year undergraduates (Darling, McWey, Howard, & Olmstead, 2007). In consequent, their university adjustment were usually poorer and their grade point average were also lower (Pritchard & Wilson, 2003). Ahmad, Noran Fuziah, Azemi, Mohd Zailani and Mohd Yusoff (2002) have conducted a study in a Malaysia Public University and found that the adjustment problems face by the first year undergraduates included course registration, understand text books written in English language, attending early lectures, health problems and financial issues such as lateness in receiving loan and scholarship or the loans and scholarship were insufficient to cater for their daily expenses.

Motivation is a factor that many researchers believed can affect academic performance of first year undergraduates, as motivation has found to be associated with effective learning and the application of new information and skills (Langley, Wambach, Brothen, & Madyun, 2004). The self-determination theory (SDT) is one of the approaches that has been employed to understand how motivational orientation affects university adjustment and the presence of stress (Clark & Schroth, 2010).

According to SDT, motivation includes three dimensions: intrinsic motivation, extrinsic motivation, and amotivation (Barkoukis, Tsorbatzoudis, Grouios, & Sideridis, 2008). Intrinsic motivation (IM) refers to engaging in an activity because it is enjoyable to do so (Fairchild, Horst, Finney, & Barron, 2005). IM has 3 categories: intrinsic motivation to know (IMTK), intrinsic motivation toward accomplishment (IMTA) and intrinsic motivation to experience stimulation (IMTES). IMTK refers to the motivation that emphasizes on the satisfaction one gains from the exploring or learning process. IMTA can be defined as the activity that one engages in for the satisfaction experienced when trying to excel, to reach a new standard, or to create something new. Lastly, IMTES occurs when individuals engage in activities for the purpose of gaining stimulating sensations such as the fun and excitement that comes from their involvement in the activity (Barkoukis et al., 2008).

In contrast, extrinsic motivation (EM) refers to motivation of performing an activity for the purpose of gaining a reward, avoiding guilt, or gaining approval (Deci, Ryan, & Williams, 1996). There are three types of EM: extrinsic motivation external regulation (EMER), extrinsic motivation introjected regulation (EMIN), and extrinsic motivation identified regulation (EMID) (Petersen, Louw, & Dumont, 2009). EMER is the least self-determined type of EM whereas EMID is the most self-determined type of EM, and EMIN occurs while the behavior is controlled by the desire to avoid certain feelings (Clark & Schroth, 2010). Amotivation (AM) refers to the absence of motivation (Vallerand, Fortier, & Guay, 1997).

Intrinsic motivation, extrinsic motivation, and amotivation with different levels of self-determination are located at different points along a continuum, which is known as the self-determination continuum (Barkoukis et al., 2008). In this continuum, amotivation represents the lower level of self-determined behavior, extrinsic motivation represents intermediate levels of self-determined behavior, and intrinsic motivation represents the higher level of self-determined behavior. High self-determined types of behavior are those that are personally meaningful, endorsed by the self, and performed by own choice. In contrast, the low self-determined types of behavior are perceived as controlled by internal or external pressures such as avoiding punishment or feelings of guilt and are considered being less internalized and less autonomous because they are not done by their own choice (Patry, Blanchard, & Mask, 2007).

Studies have supported the significant relation between motivational orientation and university adjustment, such as intrinsic motivation has been found to be positively associated with the quality of learning, lower dropout rates and greater academic persistence (Hardre & Reeve, 2003). In contrast, extrinsic motivation and amotivation have been shown to be associated with impaired learning, and poorer academic performance and educational outcomes (Benware & Deci, 1984).

Aims of the study

As of the last 10 years, students that enrolled in Malaysia's higher institutions had increased from 628,479 (in the year 2001) to 1,134,134 (in the year 2010) (Ministry of Higher Education, 2012). Therefore, it is expected that the numbers of first year undergraduates who face the problems of adjustments are also increasing. Due to the high academic expectations among Asian parents and cultural beliefs on the values of education (Hao & Bonstead-Bruns, 1998), it is important to find out strategies to assist the university adjustment among the first year undergraduates. In this study, we aimed to explore the relationship between the motivational orientations, perceived stress and university adjustment among the first year undergraduates in Malaysia.

Method

Respondents

A total of 200 first year undergraduates participated in this survey. The respondents were 96 male and 104 female. Their ages range from 18 to 24 years (M = 20.01, SD = 1.22).

Apparatus

The questionnaire contains 4 sections, which are demographic information, academic motivation scale (AMS), college adaptation questionnaire (CAQ) and perceived stress scale (PSS).

Demographic information

The respondents need to fill in their demographic information on the last page of the questionnaire, such as their age (years) and gender (male or female).

Academic motivation scale (AMS)

The academic motivation scale (AMS) was used to assess the student's motivation to succeed at university. The AMS is a 28-item measure of motivation based on the theoretical model of motivation. The AMS is composed of seven sub-scales. Three sub-scales assess types of IM (IMTK, IMTA, and IMTES), and three subscales assess types of EM (EMER, EMIN, and EMID). One sub-scale assesses AM. The respondents are asked to rate according to what extent each of the items described them appropriately on a 7-point Likert scale (1 = does not correspond at all to 7 = corresponds exactly). The AMS is scored for each of the seven sub-scales with higher scores indicating greater intrinsic, extrinsic, and amotivation. Previous studies using the AMS reveal good internal (r= 0.81) and test-retest reliability (r = 0.79) (Vallerand et al., 1992). Our results of Cronbach-alpha also found acceptable internal consistencies among these subscales, which are 0.68 (IMTK), 0.71 (IMTA), 0.71 (ITMES), 0.71 (EMER), 0.66 (EMIN), 0.62 (EMID) and 0.64 (AMS).

College adaptation questionnaire (CAQ)

The college adaptation questionnaire (CAQ) was used to assess how well students have adjusted to university life. The CAQ is a self-report instrument which consists of 18 statements scored on a 7-point scale. The respondents are asked to rate as to what extent each item described them appropriately on a 7-point scale (1 = not applicable to 7 = very applicable). Ten statements reflect poor adjustment whereas eight statements reflect good adjustment. The CAQ scores are obtained by reversing the scores on the 10 poor adjustment items and then summing across all 18 items. A higher score indicates better university adjustment. Previous studies have reported good reliability of CAQ (r = 0.83) (Rooijen, 1986). The Cronbach-alpha of this study is 0.83.

Perceived stress scale (PSS)

The PSS is a self-report measure that assesses global stress or the degree to which life situations are appraised as overwhelming, unpredictable, and uncontrollable (Cohen, Kamarck, & Mermelstein, 1983). Cohen et al. (1983) provided three versions of the PSS which are 4, 10, and 14 items, with comparable reliability for the various versions. The abbreviated four items inventory (PSS-4) was used in this study. This short version is made up of the items that correlated most highly with the original scale and has been judged to be a useful measure of perceived stress for situations requiring a short scale. The PSS-4 consists of four items rated on a 5-point Likert scale. Respondents are asked to respond to each question on a 5-point Likert scale (0 = never to 4 = very often), indicating how often they have felt or thought a certain way within the past month. The PSS scores are obtained by reversing the scores on the two positive items. Previous studies indicate the alpha reliability estimate for this scale was 0.72, and the test–retest reliability over a two months interval has been reported as 0.55. The Cronbach-alpha in this study was found to be 0.69.

Procedure

This study was conducted by using a cross-sectional questionnaire design. Purposive sampling was used in the selection of the respondents that only first year students were invited. After getting approvals from lecturers to enter the classrooms during their year one lectures, research assistants introduced the aims of the study to students and their right of not to participate in the study. Questionnaires were then distributed to students who agreed to fill in the questionnaires. All respondents were given 15 minutes to complete the questionnaire. Research assistants then collected the completed questionnaires and thanked the respondents for their cooperation.

Results

Motivational Orientations and University Adaptation

The Pearson correlation was conducted to examine the relationship between different motivational orientations and university adjustment. University adaptation was positively correlated with all the 3 types of intrinsic motivation (IMTK, IMTA and IMTES) and one type of extrinsic motivation (EMID), but was negatively correlated with AM, all significant (see Table 1).

Table 1 - Correlations between University Adaptation, Perceived Stress and Motivational Orientations (n = 200)

	Intrinsic motivation			Extrinsic motivation			
	to know	toward accomplish-ment	to experience stimulation	identified regula-tion	Introjec-ted regula-tion	external regula-tion	Amoti-vation
University Adaptation	0.410***	0.256***	0.292***	0.363***	0.021	0.042	-0.375***
Perceived stress	-0.189**	-0.148*	-0.117	0.179*	0.126	-0.144*	0.196**

Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Motivational Orientations and Perceived Stress

Again, the *Pearson correlation* was conducted to examine the relationship between different motivational orientations and perceived stress. Perceived stress was negatively correlated with IMTK, IMTA and EMER, but positively correlated with EMID and AM (see Table 1).

University Adaptation and Perceived Stress

The results of *Pearson correlation* found a negative correlation between university adaptation and perceived stress, r (198) = -0.405, $p < 0.001$. Those who reported better university adaptation have lower perceived stress.

Discussion

This study is aimed to examine the relationships between different motivational orientations, perceived stress and university adjustment among the first year undergraduates. The results found significant positive relationships between levels of self-determination with university adjustment but negative relationships with perceived stress. These findings are similar to those of previous studies, which found that most self-determined types of motivation will lead to positive and adaptive outcomes, and the less self-determined types of motivation will lead to negative and maladaptive outcomes (Vallerand, Pelletier, & Koestner, 2008).

These results are not surprising, since studies have shown motivational orientation is one of the factors that can explain a student's adjustment in university. For instance, IMTK, IMTA, IMTES, and EMID have been found to be linked with most adaptive outcomes and better adjustment in university (Petersen, Louw, Dumont, & Malope, 2010). In contrast, the less self-determined types of motivation such as EMIN, EMER, and AM are either unrelated or negatively related to adaptive outcomes and poorer adjustment in university (Wang & Liu, 2008). Importantly, our present findings suggested that these relationships are rather robust and are not limited in western cultures.

The positively relationships between EMID and both the CAQ and perceived stress are unexpected, which suggested that those high in EMID can adjust better but also perceived more stress. Park et al. (2012) argued that this could be because students high in EMID would be affected by both the extrinsic and intrinsic factors so their perceived stress were also increased.

In conclusion, this study has shown the importance of motivational orientations in university adjustment and perceived stress among the first year undergraduates in Malaysia. Based on the findings of this study, programs could be designed to make the public aware about the importance of intrinsic motivation in the choice of higher education. School counselors may need to use aptitude tests for those students who are going to graduate from their secondary schools so that they are more able choose a subject that suits their own interests. Parents should allow their children to select the subjects their own rather than what the parents thought is good for them. Nonetheless, due to the difficulties of recruiting participants from more universities, the generalization of the results should be cautions. Future studies may consider to recruit more first-year undergraduates from more universities, so that more comparisons can be made, such as the differences between those who study at private and public universities.

References

Ahmad, K., Fauziah, N., Azemi, Y., Shaari, M., & Zailani, M. Y. (2002). Adjustment to College Life and Academic Performance among Universiti Utara Malaysia Students. *Unpublished Manuscript. Universiti Utara Malaysia, Kedah, Malaysia*.

Ariffin, A., Ahmad, A. H., Ahmad, M. S., & Ibrahim, M. A. (2008). Determining decision-making styles and demographic differences in selecting higher education services among Malaysian. *International Journal of Business and Society*, *9*(1), 12–36.

Arnett, J. J. (2004). *Emerging Adulthood:The Winding Road from the Late Teens through the Twenties*. Oxford University Press.

Baker, R. W., & Siryk, B. (1984). Measuring adjustment to college. *Journal of Counseling Psychology*, *31*(2), 179–189. http://doi.org/10.1037/0022-0167.31.2.179

Barkoukis, V., Tsorbatzoudis, H., Grouios, G., & Sideridis, G. (2008). The assessment of intrinsic and extrinsic motivation and amotivation: validity and reliability of the Greek version of the academic motivation scale. *Assessment in Education: Principles, Policy & Practice*, *15*(1), 39–55. http://doi.org/10.1080/09695940701876128

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, *21*(4), 755–765. http://doi.org/10.3102/00028312021004755

Brooks, J. H., & DuBois, D. L. (1995). Individual and environmental predictors of adjustment during the first year of college. *Journal of College Student Development*, *36*(4), 347–360.

Clark, M., & Schroth, C. A. (2010). Examining relationships between academic motivation and personality among college students. *Learning and Individual Differences*, *20*(1), 19–24. http://doi.org/10.1016/j.lindif.2009.10.002

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. http://doi.org/10.2307/2136404

Colman, A. M. (2001). *Oxford dictionary of psychology*. Ne w York: Oxford Universit y Press.

Compas, B. E., Slavin, L. A., Wagner, B. M., & Vannatta, K. (1986). Relationship of life events and social support with psychological dysfunction among adolescents. *Journal of Youth and Adolescence*, *15*(3), 205–221.

Conti, R. (2000). College goals: do self-determined and carefully considered goals predict intrinsic motivation, academic performance, and adjustment during the first semester? *Social Psychology of Education*, *4*(2), 189–211. http://doi.org/10.1002/smi.1139

Darling, C. A., McWey, L. M., Howard, S. N., & Olmstead, S. B. (2007). College student stress: the influence of interpersonal relationships on sense of coherence. *Stress and Health*, *23*(4), 215–229. http://doi.org/10.1002/smi.1139

Deci, E. L., Ryan, R. M., & Williams, G. C. (1996). Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*, *8*(3), 165–183. http://doi.org/10.1016/S1041-6080(96)90013-8

Dyson, R., & Renk, K. (2006). Freshmen adaptation to university life: depressive symptoms, stress, and coping. *Journal of Clinical Psychology*, *62*(10), 1231–1244.

Elias, H., Ping, W. S., & Abdullah, M. C. (2011). Stress and academic achievement among undergraduate students in Universiti Putra Malaysia. *Procedia-Social and Behavioral Sciences*, *29*, 646–655. http://doi.org/10.1016/j.sbspro.2011.11.288

Fairchild, A. J., Horst, S. J., Finney, S. J., & Barron, K. E. (2005). Evaluating existing and new validity evidence for the Academic Motivation Scale. *Contemporary Educational Psychology*, *30*(3), 331–358. http://doi.org/10.1016/j.cedpsych.2004.11.001

Hao, L., & Bonstead-Bruns, M. (1998). Parent-child differences in educational expectations and the academic achievement of immigrant and native students. *Sociology of Education*, 175–198. http://doi.org/10.2307/2673201

Hardre, P. L., & Reeve, J. (2003). A motivational model of rural students' intentions to persist in, versus drop out of, high school. *Journal of Educational Psychology*, *95*(2), 347–356. http://doi.org/10.1037/0022-0663.95.2.347

Kerr, S., Johnson, V. K., Gans, S. E., & Krumrine, J. (2004). Predicting adjustment during the transition to college: Alexithymia, perceived stress, and psychological symptoms. *Journal of College Student Development*, *45*(6), 593–611.

Langley, S., Wambach, C., Brothen, T., & Madyun, N. (2004). Academic achievement motivation: Differences among underprepared students taking a PSI general psychology course. *Research and Teaching in Developmental Education*, *21*, 40–49.

Ministry of Higher Education. (2012). indikator_pengajian_tinggi_2009-2010.pdf. Retrieved February 26, 2012, from http://www.mohe.gov.my/web_statistik/indikator_pengajian_tinggi_2009-2010.pdf

Park, J., Chung, S., An, H., Park, S., Lee, C., Kim, S. Y., … Kim, K.-S. (2012). A structural model of stress, motivation, and academic performance in medical students. *Psychiatry Investigation*, *9*(2), 143–149. http://doi.org/10.4306/pi.2012.9.2.143

Patry, D. A., Blanchard, C. M., & Mask, L. (2007). Measuring university students' regulatory leisure coping styles: planned breathers or avoidance? *Leisure Sciences*, *29*(3), 247–265. http://doi.org/10.1080/01490400701257963

Petersen, I., Louw, J., & Dumont, K. (2009). Adjustment to university and academic performance among disadvantaged students in South Africa. *Educational Psychology*, *29*(1), 99–115. http://doi.org/10.1080/01443410802521066

Petersen, I., Louw, J., Dumont, K., & Malope, N. (2010). Adjustment to university and academic performance: brief report of a follow-up study. *Educational Psychology*, *30*(4), 369–375. http://doi.org/10.1080/01443411003659978

Pritchard, M. E., & Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of College Student Development*, *44*(1), 18–28. http://doi.org/10.1353/csd.2003.0008

Rawson, H. E., Bloomer, K., & Kendall, A. (1994). Stress, anxiety, depression, and physical illness in college students. *The Journal of Genetic Psychology*, *155*(3), 321–330. http://doi.org/10.1080/00221325.1994.9914782

Reason, R. D., Terenzini, P. T., & Domingo, R. J. (2006). First things first: developing academic competence in the first year of college. *Research in Higher Education*, *47*(2), 149–175. http://doi.org/10.1007/s11162-005-8884-4

Redhwan, A. A. N., Sami, A. R., Karim, Aj., Chan, R., & Zaleha, Mi. (2009). Stress and coping strategies among Management and Science University students: A qualitative study. *International Medical Journal*, *8*(2), 11–15.

Robotham, D., & Julian, C. (2006). Stress and the higher education student: a critical review of the literature. *Journal of Further and Higher Education*, *30*(02), 107–117. http://doi.org/10.1080/03098770600617513

Rooijen, L. (1986). Advanced students' adaptation to college. *Higher Education*, *15*(3), 197–209. http://doi.org/10.1037/0003-066X.55.1.68

Ross, S. E., Niebling, B. C., & Heckert, T. M. (1999). Sources of stress among college students. *Social Psychology*, *61*(5), 841–846.

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, *72*(5), 1161. http://doi.org/10.1037/0022-3514.72.5.1161

Vallerand, R. J., Pelletier, L. G., & Koestner, R. (2008). Reflections on self-determination theory. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 257. http://doi.org/10.1037/a0012804

Wang, C. K. J., & Liu, W. C. (2008). Teachers' motivation to teach national education in Singapore: a self-determination theory approach. *Asia Pacific Journal of Education*, *28*(4), 395–410. http://doi.org/10.1080/02188790802469052

Wintre, M. G., & Yaffe, M. (2000). First-year students' adjustment to university life as a function of relationships with parents. *Journal of Adolescent Research*, *15*(1), 9–37. http://doi.org/10.1177/0743558400151002

RELATIONSHIP BETWEEN SOCIAL PERCEPTION AND MATHEMATICAL THINKING AMONG MALAYSIAN STUDENTS IN HIGHER EDUCATION

Rommel Mahmoud Ali AlAli and **Mohamed Najib Abdul Ghafar**

Faculty of Education
Universiti Teknologi Malaysia

Abstract

This study aimed to investigate the relationship between mathematical thinking and social perception of mathematics among Malaysian students in higher education. The study sample consisted of (150) male and female students from the first year students at Universiti Teknologi Malaysia (UTM). Two instruments were developed which are mathematical thinking skills test MTT, and social perception of mathematics questionnaire SPMQ. The first instrument consisted of 24 questions and the second one consisted of 36 items. The results showed that there were statistically significant differences between the mean averages of students' scores on the dimensions of social perception of mathematics. In addition, there was a positive relationship between the difficulty of mathematics and its anxiety, and mathematical thinking.

Keywords: mathematical thinking, social perception, higher education.

Introduction

Mathematics played an important role in the development of human civilization. Human beings need mathematics in their lives, for accounts, data processing, and communication with others, problem solving, and decision-making. Mathematics occupies a privileged position between science and various fields of knowledge. Furthermore, mathematics is known as the mother of all sciences, because it is applied in various fields of life. Therefore, it is a good field to train students, develop thinking skills, and accompany the students

during their lives. Thinking and mathematical thinking are an educational necessity. Learning how to think mathematically is an extremely important issue in mathematics education. Thus, it is an essential tool for dealing with individuals in their lives in order to help them identify problems and solve those problems (Alkhateeb and Ababneh, 2011).

Social perception about the nature of mathematics, learning of mathematics and problem solving determine the strategy and methods to deal with the problems. That will lead to effective role in influencing the curriculum, and the progress of the educational process because it consists of observing and understanding of existing information and draw the conclusions. In addition, it helps the individuals to simplify and solve problems and contradictions, and help them to accept the environment (Zuhair, 2008).There is general belief for many educators that students' perceptions towards mathematics affect the extent acceptable to its concepts and experiences. Therefore, it is necessary to do everything for the development of positive perceptions among students towards mathematics as well as improving the negative perceptions towards mathematics (Aljabry, 1993). Promoting positive perceptions among students towards mathematics is still considered influential on the development of mathematical students' abilities (Lianghue and eta, 2005). Odeh (1985) indicates that most educators emphasize the importance of focusing on perception towards mathematics that is desirable and developed. However, the interest is still concentrated on the cognitive domain without the emotional domain.

However, there are negative perceptions towards mathematics. The impact of social perceptions on individuals plays an important role in their lives, because it directs the behavior to the description of phenomenon through the social framework, psychologically and biologically. The social perceptions represent thinking systems about certain subjects which differ from the opinions and attitudes. As a result of that, the individual remains connected with the community. The social perception may have an impact on mathematical thinking because it directs the learner's response towards learning mathematics (Jussim, 2012).

There are efforts to develop education and improve outcomes, but the learning outcomes in mathematics did not reach an acceptable level because of many problems like: low scientific achievement, failure to process and analyse data, giving interpretations, and solving problems (NCRD, 2007). It appears clear through the results obtained Malaysia in the International Study of Mathematics and Science TIMSS for the year 2011 showed that Malaysia's rank declined in mathematics from position 20 in 2007 to 26 in 2011, where the average score of Mathematics was 474 in 2007, fell to 440 in 2011(TIMSS, 2011). In Malaysia, most schools focus on the tests, because the results of the tests will determine the future of students in further education. Education became tends to aims on how to get high marks rather than getting knowledge furthermore, the teachers are interested in preparing students for the government examination. Also, the mismatch between Malaysian mathematics curriculums with assessment makes the students unable to use problem-solving skills and thinking (Yong, 2010). Thus, this study aimed to examine the relationship between mathematical thinking and social perception of mathematics among Malaysian students in higher education. Furthermore, the research hypotheses in this paper are: there is no significant relationship between social perception of mathematics and mathematical

thinking, and there is no significant relationship between social perception of mathematics and its dimensions and mathematical thinking.

Literature Review

The study by Sam and Yong (2006) showed that there are several obstacles to promote mathematical thinking in Malaysian schools as school culture, teachers' attitude and commitment, teachers' workload, exam-oriented culture and assessment system. The study by Hwa et al. (2009) showed that more than two thirds of students perceived that the MaTA Framework had a positive impact, helping them to improve their learning of mathematics, enabled them to sharpen their problem solving skills, promoted their thinking and let them understand mathematics better, the MaTA Framework appears capable of serving as an alternative to assessing students' thinking processes, fostering more effective and holistic mathematical thinking. Mubark (2011) in his study showed that teachers' opinions and student responses were almost the same with some changes in the order for the last two aspects (Induction and Deduction), but in term of level of difficulty, all teachers agreed that mathematical proof was the most difficult aspect among the mathematical thinking aspects, with regard to time spent in teaching the different aspects of mathematical thinking, and mathematical proof received the greatest time allocation. Zaman (2011) conducted study revealed that mathematical thinking was reasonably related to mathematics achievement with proofs, and that male student's superiority of female. Also that problem solving, induction, deduction, logical thinking, proofs and the best solution were the most important aspects of mathematical thinking with proofs as the most difficult, but generalization is the easiest these aspects. Barham and Alkhateeb (2012)in their study revealed that the following: high in deduction only, intermediate in induction, reasoning and mathematical proof, and low in the most including modeling, estimation, criticism, guessing, symbols expression and justification. Results also showed positive correlation, with a statistical significant difference between most of the mathematical thinking skills and students' achievement in mathematics. Jiar and Long (2011)in their study showed that there is significant relationship between all mathematical thinking skills and physics achievement. Kashefi et al., (2012) conducted a study to enhance mathematical thinking in the learning of multivariate calculus through the computer. The results showed to the importance of the need to use computer-based tools to enhance mathematical thinking among students, where it's considered an important strategy to help students getting ideas and design activities, and to overcome the problems.

Methodology

This study used a quantitative descriptive survey approach. The population consisted of students in Malaysian universities. Universiti Teknologi Malaysia (UTM) has been specifically chosen from Malaysia. The sample of the study consisted of first-year students for undergraduate students. Two instruments were developed in this study which are; test for mathematical thinking skills MTT and a questionnaire for social perception of mathematics SPMQ. The test is used to measure the mathematical thinking skills. It consists of six

aspects namely induction, deduction, generalization, symbolism, logical thinking, and mathematical proof. The questionnaire is used to measure students' perceptions toward mathematics. It is made up of six dimensions which are the nature of mathematics, difficulty of mathematics, the teacher's role, anxiety about mathematics, enjoyableness of mathematics, and usefulness of mathematics.

Verifying the validity of the instruments

The instruments were examined by 5 experts from UTM. Based on their opinions; the researcher modified and reformulated some items. To ensure validity and reliability the instruments were piloted with some 30 students. Responses and feedback obtained were used in modifying the final instruments. The data was analyzed using Statistical Package for Social Sciences (SPSS 15.0).

Findings

The results showed that there were statistically significant correlation between score on each item and score on sub-skill. Then calculate the matrix correlation coefficients between the sub-skills and each of them and the total score as shown in table 1.

Table 1: Matrix of correlation coefficients between the subscales and between them and total score of the MTT

Skill	Induc-tion	Deduc-tion	Symbo-lism	Logical thinking	Generali-zation	Mathema-tical proof	Total score
Induction	1	.89	.45	.37	.65	.35	.79
Deduction		1	.66	.72	.76	.53	.83
Symbolism			1	.90	.92	.96	.87
Logical thinking				1	.92	.91	.81
Generalization					1	.89	.90
Mathematical proof						1	.80

Table 1 shows that the values of correlation coefficients are relatively high, which indicates clearly that all involved in the measurement of a single concept of mathematical thinking, emphasizes the correlation of sub-scores with the total score. These results constitute an indication of the scale validity. Reliability coefficients were extracted from the responses of the sample split-half method as shown in table2. These values indicate that the test has a good degree of reliability.

Table 2: Reliability coefficients for the sub-skills the MTT

Induction	deduction	symbolism	Logical thinking	Generali-zation	Mathema-tical proof	Total score
.82	.81	.91	.90	.79	.85	.96

Table 3: Matrix of correlation coefficients between the subscales and between them and Total score of the SPMQ

Skill	Diffi-culty of math	Nature of math	Useful-ness of math	Anxiety of math	Enjoy-ment of math	Tea-cher's role	Total score
Difficulty of math	1	.59	.45	.88	.48	.71	.75
Nature of math		1	.78	.48	.66	.73	.78
Usefulness of math			1	.56	.62	.54	.76
Anxiety of math				1	.62	.63	.74
Enjoyment of math					1	.95	.81
Teacher role						1	.86

As shown in Table 3 the values of correlation coefficients are relatively high, which indicates clearly that all involved in the measurement of a single concept of social perception, emphasizes the correlation of sub-scores with the total score. These results constitute an indication of the scale validity. Reliability coefficients were extracted from the responses of the sample split-half method as seen in table4. These values indicate that the questionnaire has a good degree of reliability.

Table4: Reliability coefficients for the sub-skills the SPMQ

Difficulty of math	Nature of math	Usefulness of math	Anxiety of math	Enjoyment of math	Tea-cher 's role	Total scale
.75	.83	.78	.67	.75	.85	.86

Discussion of the results

Testing the first hypothesis that *there is no significant relationship between social perception of mathematics and mathematical thinking.*

Linear Regression was used between students' scores on both social perception scale and mathematical thinking test. Table 5 shows the correlation coefficient between the social perception towards mathematics and mathematical thinking among first year students. The correlation coefficient of 0.45 indicates that there is a positive relationship between social perception towards mathematics and mathematical thinking.

Table 5: show the correlation coefficient between the social perception towards mathematics and mathematical thinking

	social perception towards mathematics	Mathematical thinking
correlation coefficient between the social perception towards mathematics and mathematical thinking	1 0.45	0.45 1
Statistical significance	0.0	0.0

The results also show R2 equal 0.15 which indicates the ability of social perceptions towards mathematics to predict in mathematical thinking. ANOVA through F-test result shows the ability of social perceptions towards mathematics to predict of mathematical thinking is statistically acceptable, where F equal 16.10. Also the values of α and β where α equal 3.35 and β equal 0.14which are indicative of the form of, the predict equation. Mathematical thinking = 3.35 + 0.14X social perceptions towards mathematics.

To test the second hypothesis; *there is no significant relationship between social perception of mathematics and its dimensions and mathematical thinking*.

Multiple regressions were used between the students' scores on both social perception scale and mathematical thinking test. Table 6 shows the correlation coefficient between the dimensions of social perception towards mathematics and mathematical thinking among first year students. The correlation coefficient of 0.45 indicates that the relationship between the social perception towards mathematics and mathematical thinking was positive.

Table 6: show the correlation coefficient between the social perception towards mathematics and mathematical thinking

correlation coefficient	Difficulty of math	Nature of math	Usefulness of math	Anxiety of math	Enjoyment of math	Teacher role	Total score
mathematical thinking	0.08	0.35	0.12	0.45	0.10	0.18	0.45
Statistical significance	0.27	0.0	0.20	0.0	0.17	0,0	0.00

To determine the extent of contribution of the associated variables with social perceptions towards mathematics in prediction of mathematical thinking, Regression Multiple Stepwise was used.

Table7: The results of multiple regression analysis for mathematical thinking on the predicted variables among study sample

Predicted		R	R2	Increase in R2	F	Statistical significance
Mathematical thinking	Difficulty of math	0.71	0.49	0.05	15.37	0.0
	Nature of math	0.69	0.46	0.07	9.5	0.09
	Usefulness of math	-	-	-	-	-
	Anxiety of math	0.65	0.42	0.11	12.13	0.0
	Enjoyment of math	-	-	-	-	-
	Teacher role	-	-	-	-	-
Total		-	-	0.23	-	-

Table 7 shows the variables included in the regression equation and how the 0.23 variance in mathematical thinking among study sample was obtained. Three variables explained the variance. The value of the change in the square correlation coefficients were statistically significant for difficulty of math, but were not statistically significant for the nature of math and anxiety of math. These results confirmed that these three variables were the strongest variables (dimensions) linked to mathematical thinking. The difficulty of mathematics explained a variance more than the other two variables, it varies significantly among students. The difficulty of math and anxiety of math determined mathematical thinking. Because of the overlapping superiority factors and fear of failure, and the preoccupation with the level of performance on the test of the students due to their awareness that mathematics is a difficult subject and needs help to learn it. These results reveal that addressing mathematical thinking can be dealt with in different ways and may increase the learning of math and its mastery as well as allowing good performance in tests. However, the results may change the social perception of students towards mathematics, especially in their perception of the difficulty of math and their anxiety of math.

Conclusions

This paper aimed to examine the relationship between mathematical thinking and social perception of mathematics at a Malaysia higher education institution. The results showed that there was a positive relationship between the difficulty and anxiety of mathematics and mathematical thinking. In general, the relationship between the social perception towards mathematics and mathematical thinking is positive. In addition, it also shows the ability of social perceptions of mathematics to predict mathematical thinking. Also, the following dimensions nature of math, anxiety of math and difficulty of math were the strongest dimensions related to mathematical thinking. Finally, there were statistically significant differences between the mean averages of students' scores on the dimensions of social perception of mathematics.

Recommendations

Educators should focus on mathematical thinking skills in the curriculum and in teaching, taking the social perception of students toward mathematics when authoring books in mathematics teaching processes and in the preference of students and desires of the process of teaching.

References

Abed, E. & Abu Zeinah, F. (2012*).*Development of the Ability of Mathematical Thinking and Its Relation to Learning Style of Jordanian Students in Grades 8 to 10. *Journal of An-Najah University for Research (Humanities)*.Vol.26 (8).

Aljabry, N. (1993). The relationship between students' attitudes in the secondary and junior high, towards science and their personalities traits. *Journal of Social Sciences*.Vol.2(14), 92-115.

Alkhateeb, M. &Ababneh, A. (2011). The Mathematical Thinking and its Relationship with Students' Attitudes and Achievement: A Study on Seventh Grade Students in Mathematics in Jordan. *Journal of Educational and Psychological Sciences*. Vol.12, No.1, 243-266.

Barham, A. &Alkhteeb, M. (2012).Students' Levels in the Mathematical Thinking Skills for Classroom Teachers at the Hashemite University and its relation to their Achievement in Mathematics. *Educational Journal*, 26 (103), 277-312.

Jiar, Y. K., & Long, C. Y. (2011).*Mathematical thinking and physics achievement of secondary school students*. Online. Retrieved on 4/05/2013.

Jussim, L. (2012). Social perception and social reality. New York. Oxford university press, Inc.

Kashefi, H., Ismail, Z., Yusof, Y. M., &Rahman, R. A. (2012). Fostering Mathematical Thinking in the Learning of Multivariable Calculus Through Computer-Based Tools. *Procedia-Social and Behavioral Sciences*, *46*, 5534-5540.

Mubark, M. (2011). Mathematical Thinking: Teachers Perceptions and Students Performance. *Canadian Social Science,* 7(5), 176-181.

National Council of Educational Research and Training (NCERT). (2006). *Teaching of Mathematics*. (First edition), Sri Aurobindo Marg. New Delhi. National Council of Educational Research and Training.

National Council of Teachers of Mathematics (NCTM). (2000). *Principle and Standards for School Mathematics*. Reston, Va: NCTM.

Odeh, A. (1985). *Measurement and Evaluation in the teaching process*. Irbid, Dar AlAml.
Schoenberger, K. M. and Liming, L. A. (2001).*Improving Students` Mathematical Thinking Skills through Improved Use of Mathematics Vocabulary and Numerical Operations*. Eric, ED(455120).

TIMSS. (2011). *TIMSS 2011 International Results in Mathematics*. International Study Center, Lynch School of Education, MA, USA. Boston College Chestnut Hill.

Yong, H. (2010). *Development, Usability and Practicality of a Mathematical Thinking Assessment Framework.* Thesis submitted in fulfillment of the requirement for the degree of Doctor of Philosophy.

Zaman, D. A. (2011).*Relationship between mathematical thinking and achievement in mathematics among secondary school students of North West Frontier Province, Pakistan*. Doctoral Thesis. Islamabad: International Islamic University.

Zuhair, B. (2008). *Social perception of the phenomenon of suicide among university student.* Unpublished PhD thesis.The Université Mentouri Constantine.

A CASE STUDY APPROACH TO UNDERSTANDING TOTAL SURVEY ERROR AND INSTITUTIONAL RESEARCH

Sonia Whiteley

Research Strategy, Social Research Centre
and
Australian Centre for Applied Social Research Methods
Australian National University

Abstract

Total Survey Error (TSE) is a component of Total Survey Quality (TSQ) that provides a framework for assessing the extent to which a survey is 'fit-for-purpose'. While TSQ looks at a number of dimensions, such as relevance, credibility and accessibility, TSE has an operational focus on accuracy and minimising errors. Mitigating survey error involves finding a balance between achieving a survey with minimal error and a survey that is affordable.

TSE was used to provide a conceptual framework for evaluating the design of the University Experience Survey (UES) and offers a structured approach to making decisions about changing and enhancing the UES to support continuous improvement. The implications of TSE for institutional researchers that are seeking to improve the conduct of their survey research and their dataare discussed using the UES as a case study.

Total Survey Error

> *"Total Survey Error refers to the accumulation of all errors that may arise in the design, collection, processing, and analysis of survey data. A survey error is defined as the deviation of a survey response from its underlying true value." (**Biemer, 2010**)*

Total Survey Error (TSE) is the part of the Total Survey Quality (TSQ) (Biemer & Lyberg, 2003) concept that focuses primarily on the operational aspects of the survey process. A TSE approach identifies key potential sources of error in the design, collection, processing and analysis of survey data. It also provides a framework for optimising survey quality within given design and budget parameters. While TSE has been criticised for being an intellectual paradigm rather than an explanatory, statistical model of survey error (Groves & Lyberg, 2010), it does provide researchers with a strong foundation to assess, reflect on and improve their research practice.

TSE consists of sampling error, usually referred to as errors of representation, and non-sampling error or errors of measurement. Errors of representation occur as part of the sample specification and the selection of the cases from the sample frame. Non-sampling error, or errors of measurement, is broader, encompassing systematic and random errors(McNabb, 2014).

Figure 1 on the following page summarises the errors of representation and measurement that can occur at each state of the survey research cycle. The main sources of error affecting survey accuracy include sampling frame errors and omissions (e.g. gaps, biases, inaccuracies in the sampling frame), sampling errors (e.g. biases in the respondent selection routine or sub-sampling routines), measurement errors (e.g. questionnaire design errors, interviewer errors, respondents errors), non-response errors (e.g. both unit-level and item-level non-response) and data processing errors (e.g. errors in data editing, coding, weighting or the creation of data files or tables) (Biemer, 2010) (Biemer & Lyberg, 2003) (Blausis & Thiessen, 2012) (Groves, Crouper, Lepkowski, Singer, & Torangeau, 2009).

Mitigating survey error involves achieving a balance between a survey that meets the desired quality standards and a survey that is cost effective. This need for balance is particularly evident where data collection is on a large scale and requires specialist skills or infrastructure. The resulting tension between quality and expense has the potential to affect all of the components of the survey cycle. Researchers also need to be mindful that fixing one aspect of survey error can take budget away from other mitigation activities (Blausis & Thiessen, 2012). In addition, addressing one source of error can often inadvertently increase another source of error (Hillygus, 2011). For example, increasing response rates could decrease representativeness.

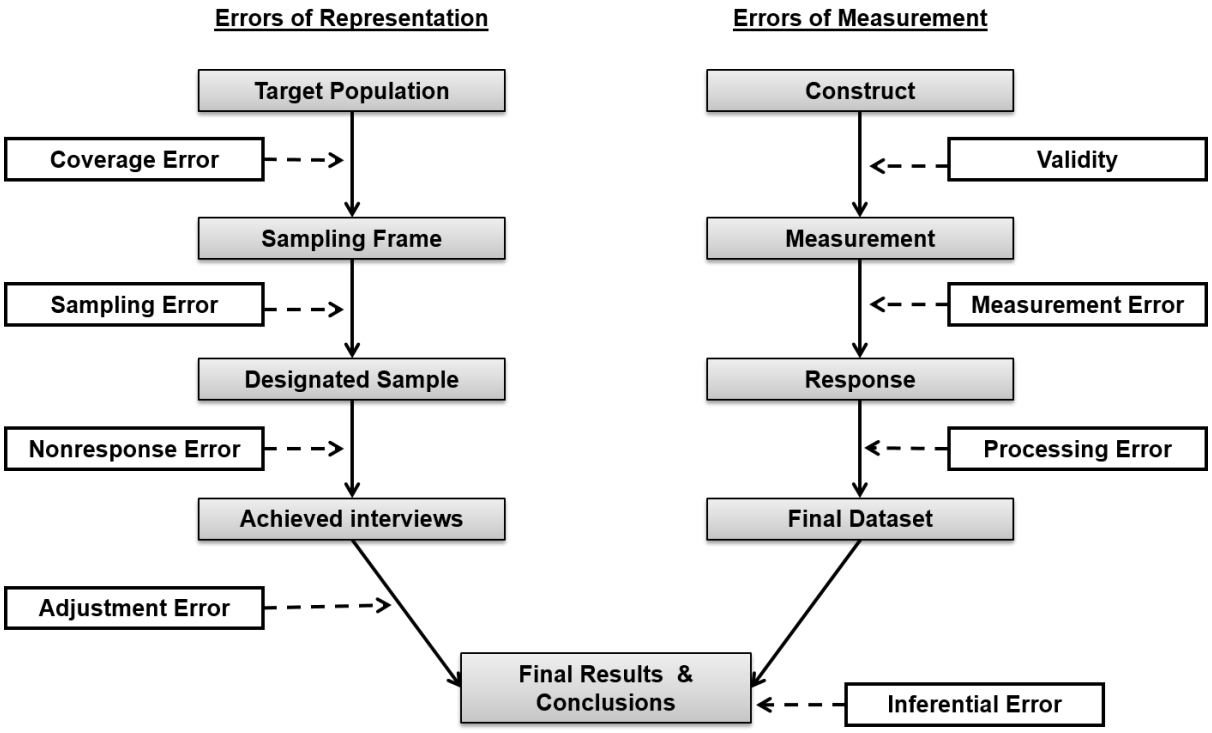


Figure 1: Total Survey Error framework in the context of the research cycle.

Adapted by (Lavrakas & Pennay, 2014) from (Groves, Crouper, Lepkowski, Singer, & Torangeau, 2009).

The intersecting and dependent nature of survey error and its mitigation means that it is essential for institutional researchers to look at errors of representation and errors of measurement from an integrated perspective rather than focusing on and remediating individual components of survey error in isolation.

Institutional research and the Total Survey Error framework

Neither a TSQ nor a TSE approach has been readily embraced by the institutional research community. This is not to say that Survey research has not been conducted in an unsystematic or inappropriate way within institutions. However, there has been the absence of an overarching quality assurance and decision making framework. Feedback obtained from institutional surveys provides key sources of data that contribute to organisational intelligence and the 'success' of the educational experience (Borden, Massa, & Milam, 2001).Therefore, it is vitally important that any criticisms regarding the robustness of survey data be addressed so that this information is on an equal quality footing with administrative data.

Liu(2010) presented a detailed conceptual strategy based on TSE for use by institutional researchers undertaking surveys. This framework addressed a useful gap in the literature

by linking institutional research practice with a means of holistically investigating and understanding survey data quality issues. However, there is little evidence that the proposed framework has been widely trialed or adopted within an institutional research context.

In the small number of instances where TSE has been used by institutions, the integrated approach to looking at a broad range of survey errors has yielded valuable insights. For example, a common measure of survey quality used by institutional researchers is the response rate. Crow, Johnson & Hanneman (2011) found that when the response rate to a survey of recent graduates was increased by using a multi-mode approach to data collection (phone, email and hard-copy form), representativeness was improved for some demographic variables but weakened for others. The variables with decreased representativeness weren't critical for the key research questions. Nevertheless, it demonstrates the subtle way that something as apparent as a higher response rate can contribute to another source of survey error.

The University Experience Survey (UES)

A consortium commissioned by the Australian Department of Education, Employment and Workplace Relations (DEEWR) designed the UES in 2011.The UES consisted of a survey instrument, the University Experience Questionnaire (UEQ), and a survey methodology (Radloff, Coates, James, & Krause, 2011). It was primarily created to measure levels of engagement and satisfaction of current first and final year undergraduate students at Australian universities. The instrument and survey approach was refined in 2012 by the same consortium.

In 2013 and 2014, Graduate Careers Australia and the Social Research Centre assumed responsibility for contributing to the continuous improvement of the execution of the UES. The UES is currently the largest survey of higher education students in Australia with 40 universities and more than 100,000 students participating on an annual basis.

A further dimension to consider in relation to the UES was presented in mid-2014 with the introduction of the Quality Indicators for Learning and Teaching (QILT). The federal budget measure committed to a survey research program aimed at collecting student feedback from undergraduate students, graduates and employers of graduates. Since the UES will form the first 'survey plank', supplemented by the Graduate Outcomes Survey (GOS) and the Employer Satisfaction Survey(ESS), it is essential to ensure that it is as robust and error free as possible prior to the introduction of the new survey elements.

A TSE issue checklist for the UES

The approach to conducting the UES for the 2013 and 2014 cycles was based on a careful consideration of potential sources of survey error tempered by an appreciation of the compressed timeline for both cycles of data collection. Therefore, it was important to assess and prioritise areas for action and improvement. TSE was used to provide a:

- theoretical and conceptual framework for evaluating the design of the UES,

- structured approach to making decisions about modifying the UES to support continuous improvement,
- method for determining a cost-effective optimal research design, and a
- means to challenge accepted paradigms regarding response rate as the primary indicator of a 'good' survey.

The TSE issue checklist on the following page was developed to identify UES specific issues that had the potential to impact negatively on data quality. The checklist was not intended to be exhaustive but was used primarily to summarise key feedback and observations made in relation to the UES during the 2012 implementation and in the lead up to the 2013 collection. Individual TSE components for errors of representation and errors of measurement are listed in the first column. The main issues for mitigation were summarised and an assessment was made regarding the extent to which the likely impact on data quality was high, medium or low.

Table 1 - TSE issue checklist identifying errors and potential impact on the UES data quality

	TSE checklist questions	**Issues**	**Potential impact on data quality**
Errors of representation			
Coverage error (under coverage and over coverage)	• How has the in-scope population been defined? • Do the specifications of the sampling frame match the population? • Are there ineligible or duplicated cases in the sampling frame?	In-scope population inaccurately or poorly defined. Sample frame may not be representative of the undergraduate population. Ineligible cases sampled.	High

	TSE checklist questions	**Issues**	**Potential impact on data quality**
Sampling error	• Is the sample size appropriate? • Is the margin of error greater than expected?	Sample size inadequate. Data not sufficiently precise for analytic or reporting purposes.	High
Non-response error	• What is the survey level non-response? • Are there any population sub-groups that do not respond to the survey? • What is the item level non-response?	High rates of survey non-response could result in non-response bias. Population sub-groups under represented. High rates of item level non-response could result in non-response bias.	High
Adjustment error	• Is the weighting schema appropriate?	Weighted data may not accurately represent the population.	Low
Errors of measurement			
Validity	• Is the instrument valid? • Is the instrument reliable?	The instrument does not measure the desired concepts or does not measure them consistently.	Low

	TSE checklist questions	**Issues**	**Potential impact on data quality**
Measurement error	• Is the questionnaire well designed? • Will interviewers unintentionally or intentionally provide incorrect information?	Poor design leading to inaccurate or incomplete responses or answers that are not relevant to the desired concepts. Interviewers may unintentionally cause respondents to change or modify their responses. Keying errors result from interviewer data input.	Medium
Processing error	• Is the analytic unit appropriately defined? • How will the data be cleaned? • Will the data be coded appropriately?	Inaccurate definition of the analytic unit. Inadequate validation checks of outputs. Coding errors or inconsistent coding of open-ended responses.	Medium
Inferential error	• Will the data be analysed and interpret correctly?	Incorrect analytic techniques used. Inaccurate inferences made.	Low

Attempting to address all of the potential survey errors during a single cycle of data collection is extremely costly and would make it difficult to determine which mitigation strategy was effective. For the 2013 UES collection, the main focus ws on reducing the errors of representation that were regarded as having the greatest impact on data quality:

coverage error and non-response error. In 2014, the error mitigation strategies shifted to address sampling error as well as retaining an emphasis on non-response error. Each of the errors of representation and the errors of measurement that were considered and actioned are discussed in detail in the following sections. Adjustment and inferential errors have not been included since their low risk rating means that they were not part of the 2013 or 2014 remediation program.

Errors of representation

An initial analysis of the errors of representation relevant to the UES suggested that coverage, sampling and non-response errors had the potential to impact substantially on data quality. Previous recommendations relevant to the UES highlighted particular concerns with the definition of the in-scope population (Radloff et.al 2011, 2012) and the lower than desired survey response rate.

In the 2013 UES survey cycle, strategies to address coverage errors and survey level non-response were prioritised. For 2014, attention was devoted to continuing to mitigate non-response errors while also focusing on sampling error. Coverage error, sampling error and non-response error mitigation approaches are discussed in further detail in the follwing sections. Errors of adjustment, data weighting have been omittedsince they were not considered to be a key issue for the data quality of the 2013 and 2014 UES collections.

Coverage error

Coverage error was identified as the highest risk to data quality. It was prioritised for remediation in 2013. The three main coverage issues that were addressed included:

- the sample frame not necessarily being representative of the undergraduate student population,
- an imprecise operational definition of the in-scope population, and
- the potential for ineligible cases to be included in the sample frame.

The fact that the UES was able to demonstrate features of under coverage, not being representative of the student population, and over coverage, including ineligible cases in the sample frame, was extremely unusual.

For the 2011 trial of the UES and the 2012 collection, a ‘bottom-up’ approach to creating the sampling frame was used where institutions provided extracts from their student systems to the contractor based on a narrative definition of the in-scope population. These records were cleaned, coded and formed the foundation of the sample frame.

The 2013 sample frame was based on a ‘top-down’ approach using population data from HEIMS to create the sample frames for each of the 40 universities. This approach minimised accidental bias being introduced into the sample selection process, ensured a nationally consistent approach to sampling and provided a sample frame that was representative of the student population.

In addition to the problem of under coverage resulting from a lack of representativeness, the narrative definition of the in-scope population was not necessarily interpreted consistently across institutions. As part of the 2013 error mitigation strategies, a clear definition of in-scope population was created syntactically using established data elements from HEIMS and applied to the national student population. All universities were provided with access to the syntax during the sample verification process for their institution to support a transparent and open approach to selecting in-scope students for the survey. The independent verification of the sampling frame also allowed institutions to flag ineligible cases, particularly students that had withdrawn or deferred from their studies or passed away and were unavailable for selection.

The success of the 2013 collection, as evidenced by a substantial increase in the participation rate from 20.2% to 29.3%and a small proportion of students identifying as out of scope for the survey (just 155 of the 330 772 students in the sample frame), indicates that using submission 1 data from HEIMS provides a robust foundation for the sampling strategy. The 2013 approach to managing coverage error was also used in 2014 and the same levels of improvement with respect to coverage error were also evident for this collection.

Sampling error

Issues relevant to sampling error were not specifically addressed in 2013 since there was insufficient information about the operational aspects of the 2012 collection to implement a remediation strategy. It was also potentially pointless to fix perceived sampling issues while the approach to creating the sample frame was being improved.The mitigation strategies identified to address the sampling error were linked and dependent, meaning that it was not possible to disentangle errors of sample size appropriateness from the expected margin of error.

The approach used to determine the appropriate sample size was broadly consistent across the 2012 and the 2013 UES collections. Commencing and 'final year' students were separately allocated to one of 45 Subject Areas used for reporting purposes on the department's website. The subject area and status (commencing or final year) groups were referred to as 'strata'. Using the assumptions outlined in the 2012 UES National Report, all eligible students were selected for strata with up to 1,333 students, effectively a census of those strata. With larger strata, a random sample of 1,333 students was drawn with the goal of yielding at least 200 responses. The value of 200 was derived from a desire for error bands of ±5% at a 95% level of confidence.

An analysis of this approach in 2013, suggested that it had a number of shortcomings in determining the sample size.. In general, large strata were substantially over sampled and often achieved completed interviews well in excess of the target of 200. The targets for each strata were uncapped and all students invited to take part were able to complete the survey.As a result, students from these large strata were substantially over- represented. This had the follow-on effect of increasing the gender imbalance since many of the large strata consisted of course offerings where males are traditionally under-represented such as nursing and education. Lastly, the sampling approach did not take into account the wide range of differential response rates across strata.

In 2014, required sample sizes were calculated at the strata level taking into account the number of records available for sampling and the requirement to report data at a 90% confidence level, +/- 5%. Using this approach to sample size identification, it was apparent that when the required sample size was compared with the response rates achieved in 2013, it would not be possible to achieve the required number of interviews for a substantial proportion of the strata. This was primarily due to the fact that for many institutions, only a small number of courses are offered in each subject area. When the 2014 sampling approach was applied, the majority of the strata were a census, rather than a sample and a response rate of up to 100% was required in order to meet the required level of reporting precision for an untenably large proportion of the strata.

In consultation with the department, the level of reporting precision was modified to a 90% confidence level +/- 7.5%. In almost all instances, the number of records that needed to be sampled was retained but the required response rate was lowered to a level that was more achievable. It was still the intention of the operational team to aim for a 5% confidence interval.This was used as a 'background target' with a view to making this the actual target in future implementations of the UES.

In both 2013 and 2014, the sample selection was checked against population parameters to confirm that appropriate proportions of gender, qualification, mode of attendance, broad field of education and citizenship characteristics were present in the sample.

Non-response error

Non-response was a clear area for improvement with high levels of survey level non-response, sub-group under representation and item-level non-response reported for the 2012 UES. The main concern for the 2013 implementation of the UES, was that the overall survey level non-response was unacceptably high.

To support the mitigation of survey level non-response error, collaborative working relationships were established with the survey managers at each of the universities. An appropriate incentivisation scheme was regarded as an important mechanism to encourage survey participation. For the 2013 UES, students were entered into a prize draw to win a small number of gift cards on completion of the survey. Students who did not respond to the initial email survey invitation were also sent a hard-copy invitation letter in addition to the email reminder.

These strategies to address survey level non-response were supported by an active monitoring of response rates throughout the data collection period. As a result, additional remedial action could be undertaken in a timely manner if institutions or program areas were underperforming. Similarly, sample representativeness was monitored and corrective action taken through targeted reminder emails and SMS' throughout the fieldwork period. Initial investigation suggests that for 2014, these activities were not as effective as anticipated and representativeness is one of the areas identified for attention in 2015.

Item level non-response for the 2013 UES was compared with the online component of the 2012 UEQ. Average item non-response to the 2013 survey was 1.0%, a substantial reduction from 7.7% average non-response to the 2012 online survey. Those students

who responded to the UES in relation to a second course, exhibited slightly higher levels of item level non-response, 1.6% on average, which is understandable given that the survey was substantially longer for this group.

An examination of some of the items with comparatively higher levels of non-response suggests that the wording of the questions is awkward which may be why students are declining to answer. Cognitive testing of these items or a revised response frame could further reduce the non-response to these UEQ items. The UEQ is scheduled for review prior to the next cycle to ensure that it aligns harmoniously with the Graduate Outcome Questionnaire and the Employers Satisfaction Questionnaire that are being developed and refined in 2014/15.

Errors of measurement

Errors of measurement associated with the UES were not the main priority for correction during 2013 or 2014 collections since, at worst, they were regarded as presenting medium levels of threat to data quality. Inferential error was seen to be a low risk. Therefore, no steps were explicitly taken to counter concerns about the interpretation of the results. Based on a review of the instrument prior to the 2013 fieldwork, validity was monitored but no remedial action was taken in 2013 or 2014. The following sections identify errors of measurement that were considered, particularly with respect to measurement and processing error and the mitigation strategies that were implemented.

Validity

Substantial effort was involved in the initial development and testing of the UEQ in 2011. These activities were not repeated during the first full deployment of the survey in 2012, largely due to the fact that there did not appear to be any issues of concern relating to the instrument.

During the initial survey set-up procedures associated with the 2013 UES, the department was provided with extensive feedback regarding the wording, sequencing and structure of the UEQ. This review of the instrument was undertaken from an operational perspective with a view to minimising measurement error. Due to the time constraints associated with the deployment of the survey in both 2013 and 2014, any modifications to the survey, aside from those required to collect data at the course or program level, were held over to be considered for future implementations of the UEQ. A number of higher priority improvements were made in 2013 to minimise survey error associated with the UES and it was more prudent to evaluate the efficacy of these modifications before addressing lower priority issues.

Even a cursory inspection of the instrument identifies issues with the face validity of the items, specificallywith respect to the use of technology and online study modes. As noted below in relation to measurement error, the wording of some items and response frames could be contributing to elevated item level non-response for a select number of questions. Consideration will be given to reviewing the instrument prior to the 2015 UEQ, particularly in light of the creation of the GOS and the refinement of the ESS.

Measurement error

Measurement error was identified as a lower priority for mitigation when compared with the higher priority concerns related to errors of representation. In 2013, a substantial change to the UES regarding the mode of data collection was implemented as a mitigation strategy. The previous phases of the UES had employed an online survey as the primary mode of data collection supplemented by non-response telephone follow-up activity.

Altering the data collection approach from mixed-mode to single online mode represented a substantial modification to the research design. Additional mitigation activities were not undertaken in 2013 or 2014 to refine the questionnaire.

Processing error

Potential processing error was initially identified as medium risk for the 2012 UES.This is primarily because the potential scope of the error was unknown and it was more conservative to monitor the data processing approach than to regard the error potential as low. Three potential processing errors were identified: inaccurate definition of the analytic unit, inappropriate coding of open-ended responses, and inaccurate data cleaning.

A clear, syntactic definition of the analytic unit was developed for the 2013 UES. The previous implementations of the UES used the student as the analytic unit. Starting with the 2013 collection, the analytic unit was the course, resulting in the inclusion of two records for students studying double degrees. To minimise processing errors associated with coding, only senior, experienced coders were responsible for coding activities. Detailed, structured procedures were developed to inform the data cleaning process. Ten per cent of each coder's workload was independently checked and queries that could not be resolved by the coding team were sent to the relevant institution for resolution.

To ensure that data cleaning did not contribute to processing error, strict input controls were implemented for the online survey to minimise the volume of data that required cleaning. Where data cleaning was necessary, primarily in relation to the open-ended responses, it was undertaken using detailed structured procedures. The accuracy of the cleaned data was independently verified by two members of the analytic team.

These effective data processing controls will be maintained for future UES collections.

Overall quality improvement & cost effectiveness

During 2013 and 2014, the pre-existing survey errors identified as presenting the highest risk to data quality for the UES were specifically addressed and minimised. Table 2 summarises the relevant survey errors, the key drivers of improved quality and the change in assessed risk ratings. With the exception of validity, which wasn't explicitly included in the survey error improvement program, all types of errors identified for mitigation were reduced. The greatest improvement was made in relation to coverage error.While there is still some room for further improvement, for example by tweaking the algorithm used to identify completing students, there is little value to be gained by

continuing to devote intense effort to address over coverage or under coverage of the in-scope population.

Table 2 - Key drivers of improved quality of the UES and risk ratings for survey error.

Type of error	Key drivers of improved quality	Original risk rating	Current risk rating
Coverage error	Use of HEIMS, robust definition of 'completing' students, rigorous institutional validation process	High	Low
Sampling error	Improved response rates, refined sampling strategy, fine-level quota control of study areas within institutions	High	Medium
Non-response error	Increased unit level response rates, reduced item level non-response, and stable under representation of some sub-groups.	High	Medium
Validity	-	Low	Medium
Measurement error	Use of a single mode of data collection and data input logic controls.	Medium	Low
Processing error	Clear definition of analytic unit, documented data cleaning and file preparation procedure, well-developed quality checks for coding and an effective query resolution process	Medium	Low

Sampling error, non-response error and validity still potentially present some risk to the data quality of the UES.These areas will be targeted for improvement in 2015.Given the nature of the quota groups and the requirement for substantial response rates for those institutions that have smaller enrolments but a large number of subject areas, this will continue to be a challenge for future iterations of the UES. Validity was not explicitly addressed in 2013 or 2014, but there is some evidence to suggest that it would be appropriate to review this potential contributor to survey error. As a result, the risk rating has increased for validity from low to medium.

One of the useful features of TSE is the appreciation of the desire to achieve the best survey quality outcomes within a typically constrained, budget. Essentially, the cost of implementing the error mitigation strategies for the 2013 and 2014 UES had a neutral

impact on the overall budget when compared with the 2012 cycle of data collection. The money saved by just using a single mode of data collection instead of supplementing online responses with telephone interviews was used to undertake the response maximisation activities required to reduce the unacceptably high survey level non-response. Decreased survey error in relation to coverage error, sampling error and processing error resulted from the use of established ISO procedures and some useful ideas about the way that pre-existing data and resources, such as HEIMS, could be used creatively and effectively.

It is important to acknowledge the contribution of the survey managers and fellow institutional researchers who actively supported, tested and engaged in the survey and error mitigation process. The cost of their time has not been taken into account but has undoubtedly contributed to the reduction of a range of error sources, particularly during 2013 when coverage and non-response errors were being addressed.

Conclusion

The implementation of a TSE framework using the UES as a case study, provides an example of the types of issues that arise and need to be taken into account when designing and implementing a large-scale survey program. It is possible that some institutional survey managers can become complacent when undertaking survey research. This is because they are usually executing a census of a population from a sample frame that they have direct access to via a student administration system. While this may reduce the need to focus on sampling related errors, there are still a number of errors that could be present in the overall survey design. The analysis of the threats to the quality of the UES data clearly shows that errors of representation and measurement are present in a survey program that has been rigorously tested and developed. This highlights the need for the monitoring and continuous improvement of survey designs and processes.

In general, institutional researchers planning to undertake or review large-scale survey programs could benefit from:

- using TSE as a conceptual framework to provide a basis for identifying and classifying relevant survey errors, and
- undertaking a risk assessment of the identified surveys errors in order to prioritise the errors for remediation.

A formal assessment of the survey error and the associated risks to subsequent data quality also provides the basis to evaluate whether or not planned mitigation strategies have the potential to be cost effective. For example, it may be worthwhile to implement a relatively expensive response maximisation strategy if high levels of survey non-response are likely to have a substantial impact on data quality.

Finally, having a clear understanding of the survey errors and the associated threats to data quality supports an additional approach to prioritising mitigation activities. Errors that occur earlier in the research cycle can flow through to later data collection activities. Therefore, it may be worthwhile to consider rectifying these errors, even if they are a lower priority, before moving onto 'downstream' issues. It may be the case, for example, that undertaking additional surveys in another data collection mode to improve

representativeness may not reduce the overall survey error if sampling errors have not been addressed.

TSE is obviously not a 'magic bullet' for fixing all institutional research survey quality issues but it does provide a clear foundation for evidence based decision making about the quality of survey design, implementation and the resulting data outputs. It is also a useful means for institutional researchers to reflect on and continuously improve their practice with respect to survey research.

Bibliography

Biemer, P. (2010). Total Survey Error: Design, Implementation and Evaluation. *Public Opinion Quarterly, 74*(5), 817-848.

Biemer, P., & Lyberg, L. (2003). *Introduction to Survey Quality.* USA: John Wiley & Sons.

Blausis, J., & Thiessen, V. (2012). *Assessing the Quality of Survey Data.* USA: Sage Publications.

Borden, V., Massa, T., & Milam, J. (2001). Technology and tools for institutional research. In R. (. Howard, *Institutional Research: Decision Support in Higher Education* (pp. 201-209). Tallahassee, Florida.

Crow, D., Johnson, M., & Hanneman, R. (2011). Benefits and Costs of Multi-Mode Survey of Recent College Graduates. *Survey Practice, 4*(5), 1-6. Retrieved August 8, 2014, from http://surveypractice.org/index.php/SurveyPractice/article/view/86/pdf

Groves, R., & Lyberg, L. (2010). Total Survey Error: Past, present and future. *Public Opinion Quarterly, 74*(5), 849-879.

Groves, R., M, F. F., Crouper, M., Lepkowski, J., Singer, E., & Torangeau, R. (2009). *Survey Methodology.* New Jersey: Wiley.

Hillygus, S. (2011). The Practice of Survey Research: Changes and Challenges. In A. Berinsky, *New Directions in Public Opinion* (pp. 32-51). Routledge.

Lavrakas, P., & Pennay, D. (2014, 3 21). *QPMR Webinar: Total Survey Error.* Retrieved 9 19, 2014, from The Social Research Centre: http://www.srcentre.com.au/news-events/news-and-events-item/2014/03/21/qpmr-webinar-total-survey-error

Liu, Q. (2010). A Survey Data Quality Strategy: The Institutional Research Perspective. *Association for Institutional Research*, (pp. 1-34). Atlanta, Georgia.

McNabb, D. (2014). *Nonsampling error in social surveys.* USA: Sage.

Radloff, A., Coates, H., James, R., & Krause, K. (2011). *Report on the Development of the University Experience Survey.* Canberra: Australian Government.

Radloff, A., Coates, H., Taylor, R., James, R., & Krause, K. (2012). *2012 University Experiene Survey National Report.* Canberra: Australian Government.

EFFECTS OF USING CONTRASTIVE ANALYSIS GRAMMAR INSTRUCTION TO ENHANCE WRITING ABILITY OF THAI EFL STUDENTS

Thanawan Suthiwartnarueput
Faculty of Liberal Arts, Mahidol University, Thailand

Abstract

Every language has distinctive features which are different from one another. The English language is a West Germanic language which was first spoken in early medieval England. The Thai language, however, belongs to the Tai-Kadai language family and more than half of the words were derived from Pali, Sanskrit and Old Khmer. Because they originated in greatly different regions, Thai and English undoubtedly have many different language aspects. In consideration of such a big difference, the present study explored the effects of Contrastive Analysis Grammar Instruction on the improvement in writing ability of 94 Thai EFL (English as a foreign language) undergraduate students at a public university in Bangkok. The data were collected from the pre-test, post-test and interview responses. The findings indicated a greater improvement in writing ability of the experimental group that received Contrastive Analysis Grammar Instruction than those who were taught in a traditional approach with statistical significance ($t = -8.10$, $p \leq 0.05$). According to the interview results, the students claimed that the English and Thai languages had many different features that brought some difficulty to Thai EFL students, but with Contrastive Analysis Grammar Instruction, they wrote English sentences more correctly.

Keywords: Contrastive Analysis Grammar Instruction, negative transfer, positive transfer, writing ability

Introduction

Writing is one of the productive skills needed for various fields of study and work. Typically, many occupations require good writing ability. We write memos, email, letters and notes at work. Students write essays, journal articles, reports and many other pieces of writing at school or university. Mostly, people who have good writing ability are seen more credible than those with poor writing ability. With better writing ability, they can gain success in communication. They can inform, convince, instruct or inspire other people easily. Per contra, people with poor writing ability usually have communication problems. They can mislead the audience if they write English texts ungrammatically or use wrong words. Because of these reasons, writing ability is undeniably important for academic, professional and personal success. The present study considers the importance of writing ability for academic success and uses Contrastive Analysis Grammar Instruction, teaching language by comparing and contrasting grammatical features between the English and Thai languages, to improve writing ability of Thai EFL undergraduate students.

Contrastive Analysis

During the structuralism period, linguists paid attention to examining different features of the native language and those of the foreign language since the differences between two languages could cause foreign language learning difficulty. That thought gave birth to Contrastive Analysis which originated in Central Europe before the Second World War. The theory lays great emphasis on language awareness and a connection between aspects of languages (James, 2005; Kortmann, 1996; Mair, 2005; Ridwan, 1998). Jawasi (2006) mentioned it as an explicit method to analyze the differences and similarities of two or more languages or even sub-system of language. Its purposes include making foreign language instruction more effective through presenting the differences between the native language of learners and the target language. That is to say, the learners' native language has an influence on learning a foreign language both facilitating learning (positive transfer) and causing problems (negative transfer or interference).

Contrastive Analysis bloomed in the 1960s and early 1970s. It was widely used as a method of explaining some difficult features of a target language. Contrastive Analysis supporters claimed that comparing two languages assisted the teacher in preparing L2 (second language) teaching materials and predicting learners' behaviors and difficulties. It was believed that when similarities and differences between L1 and L2 (first and second languages) were taken into consideration, the teaching progress could be more effective and useful. Such beliefs led to the birth of Contrastive Analysis Hypothesis which assumes that the main difficulties found in learning a target language are based on the linguistic structure comparison of two languages. In the United States, after Charles Fries had formulated the theory in 1945, in the late 1950s, Lado (1957) was another proponent who claims that Contrastive Analysis is a means to identify areas of difficulty for language learners. He wrote and developed the theoretical foundations of what today called the Contrastive Analysis Hypothesis in the book "Linguistics across Cultures". In the book, it is said that the language elements which are similar to the learner's mother

tongue will be easier to learn while those different elements will be difficult. In other words, language learners can gain success in learning when their native language is similar to the target language. That is called positive transfer. Schuster (1997) claims that English learners of German language or German learners of English language can acquire the target language through positive transfer because the two languages have many similarities. Additionally, König and Gast (2009) have recently stated that advanced learners can learn better from a direct comparison of their mother tongue with the target language. For this sake, course materials in contrastive linguistics have been developed and included in teacher training programs in many universities.

Considering its development, in the earlier years, Contrastive Analysis seemed to focus on the linguistic elements, like grammar and the lexicon according to James (1980), but in the 1980s and 1990s, the focus of the theory was broaden to cover various language elements, such as discourse structure, contrastive sociolinguistics, cross-cultural pragmatics, and contrastive rhetoric as well as speech communities in terms of their degrees of directness, explicitness, etc., all of which can lead to intercultural misunderstandings (Connor, 1996; Hellinger & Ammon, 1996; Wierzbicka, 1985; 1992). For example, scientific texts are differently structured in the Continental European from the Anglo-Saxon tradition. This is because the differences of text structuring vary from one language to another (House, 1996).

Although Contrastive Analysis seemed to get much attention, in the late 1970s, its popularity was reduced and no longer used much as it once did before. Many opponents believed that errors that appeared to be difficult could not be predicted with Contrastive Analysis. Such an issue led to the criticism to the Contrastive Analysis Hypothesis (Larsen, et al., 1991). Abbas (1995) found that adverbial positions of both English and Arabic can be put more than one place and that means Arabs does not have much difficulty in learning English adverbials. In addition, in the 1990s, pedagogical applications of Contrastive Analysis did not seem to gain much attention and there was a steady decrease in the number of textbooks (Sheen, 1996). Abbas (1995) asserts that Contrastive Analysis lays greatly on interference, thus blocking learners from predicting other important errors they are likely to commit. Klein (1986) found that when Turkish learners of German followed the grammatical structure of their native language, they often put the verb into the final position. So did Spanish and Italian learners even though in their own languages, verbs are not in final positions. For this reason, interference of a learner's native language is not the only factor in language learning difficulties. Klein (1986) reveals that both structural similarities and differences between two linguistic systems and actual production and comprehension are not the same things. The opponents believe that Contrastive Analysis is only a matter of the linguistic systems or structures, but acquisition is dealt with comprehension and production. In other words, a specific structure of a second language structure can be easy to comprehend but hard to use. Thus, prediction of possible language transfers should not rely on comparisons of structural properties but when learners use such properties (Klein, 1986). Another weak point is that Contrastive Analysis puts great emphasis on the role of interference as a source of errors without considering some other parameters of language acquisition like ages, the difference between mediated and natural acquisition, and the difference between

the acquisition of second and third languages (Klein, 1986). As a result, in the 1970s, Contrastive Analysis was of little use among language teachers since they believed that this method was unreliable. People who had been in support of Contrastive Analysis moved their attention to learners' varieties, error analysis or interlanguages (Corder, 1967; Nemser, 1971; Selinker, 1972).

However, Major (2001) argues that minimally different features in spelling which are abandoned can result in poor performance on related sounds, but noticeable differences are more easily perceived. Oldin (1989) gives an example of the similar spelling between English "embarrassed" (ashamed) and Spanish "embarazado" (pregnant) that can convey a wrong meaning to an Englishman. At any rate, Fisiak (1981) believes that both similarities and differences can equally cause a lot of problems in language learning. The versions of Contrastive Analysis were noticed then.

Wardhaugh (1970) asserts that Contrastive Analysis can be divided into two versions: strong and weak. The strong version sheds light on that most of the errors the students make when they learn a foreign language are resulted from negative transfer. On the other side, the weak version insists on explaining errors after the students have made (Wardhaugh, 1970). Oller and Ziahosseiny (1970) introduced a moderate version of the Contrastive Analysis Hypothesis. They claim that language patterns which are minimally different from the student's native language are harder to acquire than distinctive features. As discussed in their study about English spelling errors on the UCLA (University of California Los Angeles) placement test, it was found that foreign students whose native language employed a Roman alphabet had more trouble acquiring another roman spelling system than those whose native language had little or no relation to the Roman alphabet. Proponents of Contrastive Analysis have used different ways to promote its moderate version and those include the "surface structure" which corresponds to heard or spoken language or "deep structure" which underlies the language meaning or an abstract representation. They are not completely reliable, though. James (1980) claims that using the "surface structures" to compare the similarities in two languages has some disadvantages because categories found in two languages which are identified the same can have very different conditions of use. As shown in the following examples, Sentence 1 and Sentence 2 seem to have the same "surface structure", but they are used in different real-life situations. However, Sentence 3 has the different "surface structure," it is pragmatically equivalent to Sentence 1 though.

Sentence 1. The postman opened the door. (English)
Sentence 2. Le facteur ouvrit la porte. (French)
Sentence 3. Le facteur a ouvert la porte. (French) (James, 1980, p. 169)

Due to the inconsistencies of languages as shown above, contrastivists (Contrastive Analysis experts) pay more attention to the "deep structure" that could be a better approach for making comparisons. However, they need to take into account that similar look sentences of different languages with a common "deep structure" are not always equivalent in terms of their communicative potential. For example, the sentences like "The door was opened by the postmen" and "Le facteur a ouvert la porte" have a

common origin but they are not exactly the same in terms of their communicative functions.

As mentioned above, Contrastive Analysis is quite time-honored even though it has been rejected by some scholars. The instruction itself can bring about positive and negative transfer and that raises the question if Thai students whose native language is greatly different from English can get maximum benefit from Contrastive Analysis.

Differences between Thai and English Grammatical Features

English grammatical structures are one of the linguistic elements (grammatical structures, vocabulary items, and discourse) that interfere with writing ability of Thai EFL students because many English grammatical features do not exist in the Thai language.

A large number of studies have showed that syntactic structures in Thai obstruct the students' ability to use English structures. Bennui (2008) claims that Thai does not have "articles"; as a consequence, it is possible for Thai EFL students to omit them in their pieces of writing. For example, "Siam Hotel is comfortable and luxurious hotel." or "Sue is intelligent girl.". The sentences show the omissions of "articles" that can happen in pieces of writing of Thai EFL students. Bennui (2008) reveals that Thai grammatical structures interfere with the expletive construction (there + verb be) since in Thai, there is only the verb "mee" (mi:) to show possession. Students are likely to translate the verb "mee" in Thai into the verb "have" in English. As a result, Thai EFL students seemingly write "My house has three people" or "There has many trees in my house." Also, based on the English paragraphs written by 28 third-year English-minor students in Thailand, he found seven categories of grammar errors, such as word order of Thai structure, tenses, subject and verb agreement, infinitives, the verb "have", prepositions, and noun determiners, all of which were caused by their mother tongue's syntactic interference. Thai EFL students overused the preposition "with" instead of "at" and "to" because the preposition "with" /kaab/ has meanings like "at" and "to" (Bennui, 2008).

Apart from the stated problems above, other researchers found that Thai students had difficulty learning English grammatical features. In the paragraphs written by 100 first-year university students in Thailand, Pongpairoj (2002) found that Thai EFL students incorrectly used English prepositions. They overused the preposition "on". For example, "There are birds on the sky" "The boy is sleeping on the bed." They also had problems with using countable and uncountable nouns. For example, "She has black hairs." and "The room was full of furnitures." Pongpaitoj (2002) reported that the students made errors of relative clauses and the structure of "there + verb be" in English. The students had made run-on sentences, fragments, and incorrect word order, and used incorrect demonstrative and indefinite determiners as well. Such errors were resulted from direct translation from Thai into English, one of the factors in mother tongue's syntactic interference. Noochoochai (1978) studied the temporal aspect of Thai and English and found that Thai students used "will" and the base form of verbs for every temporal aspect in English. Pengpanich (2002) reported that Thai EFL students usually had language learning difficulty and tended to avoid using an "ing" ending and a "to" plus infinitive

form. Rungnaphawet (2005) found that Thai EFL students had problems with English adverbs. They did not understand forms, functions, and positions of English adverbs. Especially the first and middle positions of English adverbs, students lacked the ability to place adverbs in the right position because adverbs in Thai were not normally placed in those places. Khamput (2004) analyzed diaries written by Thai high school students and found three categories of linguistic errors, such as interference, interlingual, and developmental and uncategorized errors. Most common errors included omission of the subject (i.e. I think is very romantic.), using adjectives as main verbs (i.e. I happy very much.), and using Thai word order in English sentences (i.e. I buy books a few.) (Khamput, 2004).

As a matter of fact, difficulty in language use and errors are mostly from direct translation, interference, and nonexistence of some features. Likitrattanaporn (2002) reported that from the English paragraphs and essays of Thai EFL third-year university students, it is true that the students directly translated Thai into English sentences. For example, "Plan of my future will volunteer for the social example of warrior's die the war." and "I made the English homework." Bootchuy (2008) conducted a study with 41 first-year Thai graduate students at a university in Bangkok. From 41 pieces of writing and final term papers of the students, interference errors were found the most. Such errors included objects and complements; omissions of subjects, auxiliary verbs, subordinators, and transitive verbs; incorrect uses of adjectives, phrasal verbs, modifiers, quantifiers, and compound nouns; using serial verb constructions; using "verb to have" after "there"; incorrect positions of subjects and modal verbs in questions; run-on sentences; and fragmentation. Ratanakul (2000) studied verb-tense errors in Thai EFL students' pieces of writing and found that their language errors were resulted from the nonexistence of subject-verb agreement, verb tenses with morpheme addition or form change, and auxiliary verbs in the Thai language. Khaoural (2002) found that most of Thai EFL students transferred rule patterns in Thai into their English writing. In consequence, they made various errors, including grammatical errors (i.e. tenses, determiners, prepositions, and verbs), syntactic errors (i.e. contraction form, incomplete sentence structure, word order, compound sentences, and punctuation), and lexical errors (direct translation from Thai into English, spelling, overgeneralization of and using common lexical items).

The previous studies as discussed reveal that Thai students had problems with their English writing ability since it was obstructed by language interference, on the grounds that many English grammatical features are totally different from those of Thai. In consideration of such a matter, this raises a question if there could be any instructional methods to assist English language learning; the answer could be Contrastive Analysis. The present study gives place to Contrastive Analysis and has explored the effects of Contrastive Analysis Grammar Instruction on the improvement in writing ability of Thai EFL undergraduate students.

Research Questions

1. To what extent did Contrastive Analysis Grammar Instruction improve writing ability of Thai EFL students?

2. What effects did Contrastive Analysis Grammar Instruction have on writing ability?

Definitions of Terms

Positive transfer occurs when the similarities between the two languages help students learn the target language easier and faster. For example, the English and Thai languages have SVO (Subject-Verb-Object) structure and that helps Thai students to make English sentences.

Negative transfer or interference occurs when some different aspects of the target language make learning more difficult, such as the contrast of noun phrase positions of Thai and those of English.

Contrastive Analysis Grammar Instruction is a term used specifically in this study to refer to a teaching method that compares and contrasts grammatical aspects of the English and Thai languages. It includes both positive and negative transfer methods.

Methods

Participants

The participants were 94 Thai EFL undergraduate students of the 2014 academic year at a public university in Bangkok, Thailand. Their ages were between 18 and 20 years. The students were from engineering, science, nursing, and pharmacy faculties and they had been studying English for 12 years. They were classified as pre-intermediate students and placed in the English pre-intermediate course using the O-NET score or Ordinary National Education Testing (a test that students in the final year of high school take before university admission) in the English subject of more than 45 points. The course mainly focuses on the uses of English vocabulary, expressions and grammar in writing, reading, listening and speaking. Concerning the group selection, the students in this group seemed to have much difficulty understanding English grammar and that caused a lot of problems in writing, though their other language skills were fairly acceptable.

Teaching Techniques

The teaching period lasted for three weeks with six hours for each. The students were divided into two groups: experimental and controlled groups with forty seven students for each. The students in the experimental group received Contrastive Analysis Grammar Instruction, while those in the controlled group were taught in a traditional way. The students in the experimental group were taught in three steps as follows:

1. Negative Transfer

The teacher presented the chart to students. It shows the differences between the English and Thai languages. As shown, the English noun phrase begins with quantifiers or

cardinal numbers followed by adjectives and nouns; conversely the Thai noun phrase starts with nouns followed by adjectives and quantifiers or cardinal numbers.

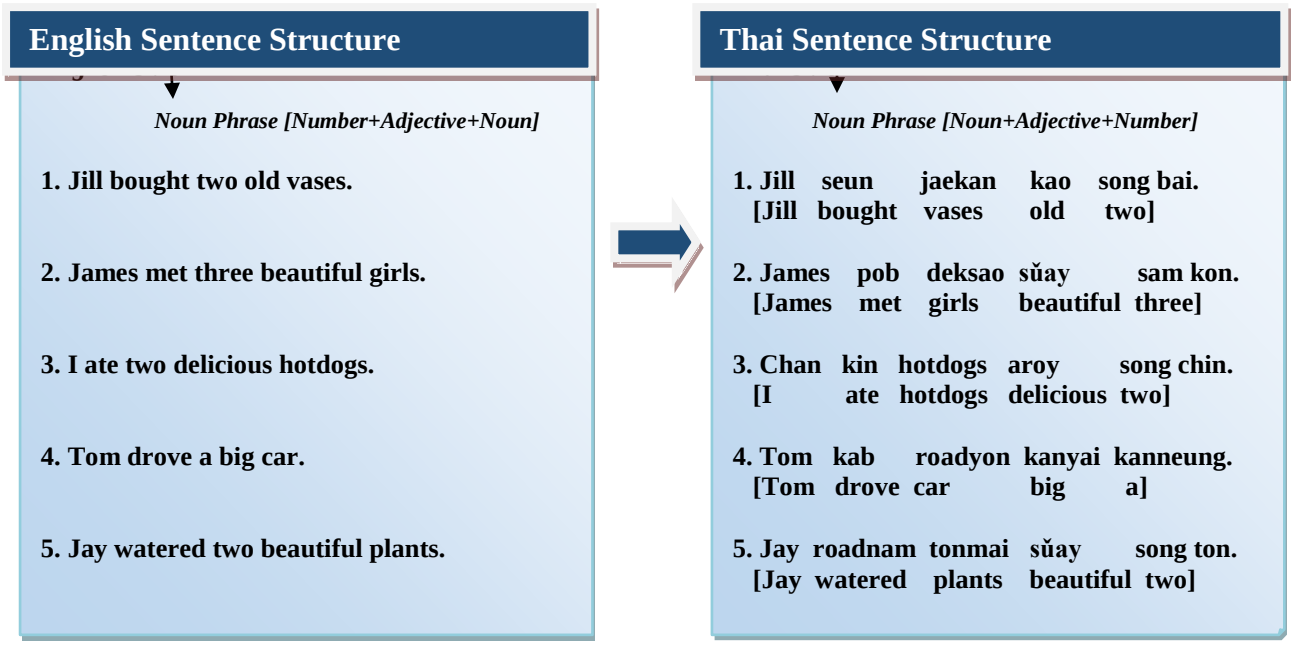
Word Order in the English Noun Phrase	Word Order in the Thai Noun Phrase
1. Five thin ducks	1. Ped pom hah tua [Ducks thin five]
2. Two beautiful girls	2. Deksao sŭay song kon [Girls beautiful two]
3. Ten white vases	3. Jaekan srikao sib bai [Vases white ten]
4. A big house	4. Baan langyai neung lang [House big a]
5. Three green bags	5. Toong srikyew sam bai [Bags green three]

2. Positive Transfer

English Sentence Structure	Thai Sentence Structure
1. Alice eats rice.	1. Alice kin kao. [Alice eats rice.]
2. Tom reads novels.	2. Tom arn niyai. [Tom reads novels.]
3. Jack studies mathematics.	3. Jack lian kamnuan. [Jack studies mathematics.]
4. Sue cooks dinner.	4. Sue tam-a-han yen. [Sue cooks dinner.]
5. Mike speaks English.	5. Mike pud pasaa aungkrit. [Mike speaks English.]

Even though a lot of negative transfer can occur and interrupt the students' learning achievement, some similarities between the two languages can assist students in learning. It is called "positive transfer". The chart above shows that both English and Thai are SVO (Subject-Verb-Object).

3. Negative and Positive Transfer



However, sentences with noun phrases can make both positive and negative transfer. As illustrated in the chart above, even though the whole sentence of both English and Thai is SVO (Subject-Verb-Object), under the object part, it can consist of noun phrases which are different between the two. The English sentence "Jill bought two old vases.," for example, the subject "Jill" is followed by the verb "bought". In the same way, the Thai sentence "Jill seun jaekan kao song bai." shows that the subject "Jill" is followed by the verb "Seun" (bought). However, the word orders in the object part of the languages are different. The English noun phrase "two old vases" is equivalent in meaning to "jaekan (vases) kao (old) song bai (two)," but they are placed in different positions.

Pre-Test and Post-Test

The topics for writing were "My Favorite Pet" for the pre-test and "The Animal in My Heart" for the post-test. For each topic, the students were assigned to write at least 100 words with the 45-minute time allotment. The total score was 50 marks graded from the criteria which included format (3 marks), organization (10 marks), content (10 marks), punctuation and mechanics (10 marks), grammar and sentence structure (10 marks), and overall impression (7 marks).

Pre-Test
"My Favorite Pet"

Write a paragraph about your favorite pet of at least 100 words. Use the right paragraph format, punctuation marks, and English grammar and sentence structure. You have 45 minutes to complete it. The total score is 50 marks.

.....
.....
.....
.....
.....
.....
.....
.....

Post-Test
"The Animal in My Heart"

Write a paragraph about the animal you like of at least 100 words. Use the right paragraph format, punctuation marks, and English grammar and sentence structure. You have 45 minutes to complete it. The total score is 50 marks.

.....
.....
.....
.....
.....
.....
.....
.....

Research Procedures

The procedures began with collecting needs of the students and it revealed that they had trouble with using subject-verb agreement, passive voice, word order in noun phrases and sentences, and verb tenses since these grammatical features are greatly different from those of the Thai language as illustrated below.

A. Subject-Verb Agreement

In English, verbs can indicate the time and their forms must agree with the number of the subject.
For example,

Past event
"Kao len tennis Mauwan." (Thai)
"He played tennis yesterday."(English)

Present event
"Kao len tennis tookwan."(Thai)
"He plays tennis everyday."(English)

The verb "len" in Thai does not require any morphemes to show the past event or to make it agree with the subject.

B. Passive voice

In English, the passive voice is formed by the verb "be" and the past participle, but in Thai, auxiliary verbs or changes to the verb are not required for making the passive.
For example,

"Rongrien knogchan tang you bon thanon nee."(Thai)
"My school is located on this road." (English)

From the above examples, there is no auxiliary verb (like "is") or any changes to the verb form (like "located") for making the passive in Thai

C. Word order in noun phrases and sentences

In English, adjectives precede a noun, but in Thai, adjectives come after a noun.

For example,

"a beautiful girl" (English)

"Bhū-ying sŭay kon neung." (Thai)

The noun "Blu-ying" (girl) comes before the adjective "sŭay" (beautiful) and the determiner "kon neung" (a) in Thai.

However, the English sentence structure is similar to the Thai sentence structure since both are SVO [Subject-Verb-Object].

For example,

"I ate pizza" (English)

"Chan kin pizza." (Thai)

D. Verb tenses

Verb tenses are clearly defined and used to indicate the time in English. The morpheme like "-ed" can be added to the verb. If they are irregular, their forms can change such as "eat", "ate", "eaten". Sometimes auxiliary verbs can be added before the main verb, such as "has broken", "has been broken", "was broken", etc. Such indicators do not exist in Thai though. The time is normally implied through temporal adverbs.

For example,

Past event

"I went to school on time yesterday."(English)

"Chan pai rongrien tong weila mauwan." (Thai)

Present event

"I go to school on time."(English)

"Chan pai rongrien tong weila."(Thai)

From the above examples, the temporal adverb "mauwan" ("yesterday" in English) indicates the past event while the form of the main verb "pai" ("go", "went" in English) is still the same in Thai.

After knowing their problems, the pre-test was administered to the students. Then the researcher randomly divided the students into two groups: experimental and controlled groups. During the experiment, the students in the experimental group were taught to use English grammar in general, mainly emphasizing subject-verb agreement, passive voice, word order in noun phrases and sentences, and verb tenses using the methods as shown in the part "Teaching Techniques". After the three-week period, the students from both groups took the post-test and were interviewed.

Data Analysis

1. The pre-test and post-test scoring data of the experimental group were compared with those of the controlled group using Independent Samples T-Test to compare the writing ability between the groups.

2. The pre-test scoring data of both the experimental and controlled groups were compared with the post-test scoring data using Paired Samples T-Test to find out their improvement after the treatment.
3. The pre-test and post-test were graded by two raters and then the scoring results were analyzed to ascertain the reliability of the scoring process. The inter-rater reliability co-efficient was 0.96 for the pre-test and 0.94 for the post-test.
4. The interview data were transcribed and analyzed by content analysis. The data retrieved were used to support the results of the pre-test and post-test.

Results

Improvement of Students' Writing Ability

As shown in Table 1, the students in the experimental group showed a great improvement in writing ability after they had learned English grammar from Contrastive Analysis Grammar Instruction. As shown in the table, their mean score rose from 36.30 in the pre-test to 43.85 in the post-test, along with the t-value to be -9.02. That is statistically significant at $\alpha \leq 0.05$.

Table 1: Comparison between the Pre-Test and Post-Test Writing Scores of the Students in the Experimental Group

Test	N	Mean	Std. Deviation	t	Sig.(2-tailed)
Pre-Test	47	36.30	3.68	-9.02	.000
Post-Test	47	43.85	4.30		

Considering the writing ability of the controlled group, however, they seemed not to show much improvement with a traditional technique. As shown in Table 2, their mean score did not raise much from 35.94 in the pre-test to 37.21 in the post-test. The t-value was -1.53, which is not statistically significant at $\alpha \leq 0.05$.

Table 2: Comparison between the Pre-Test and Post-Test Writing Scores of the Students in the Controlled Group

Test	N	Mean	Std. Deviation	t	Sig.(2-tailed)
Pre-Test	47	35.94	4.03	-1.53	.133
Post-Test	47	37.21	3.62		

Table 1 and Table 2 above reveal that the students who received Contrastive Analysis Grammar Instruction received a greater advantage in learning than those who received a traditional treatment, though in the beginning before both groups received the treatments, their writing ability was not greatly different. As shown in Table 3, the mean score of the

controlled and experimental groups were slightly different, along with the t-value to be -.45, which is not statistically significant at $\alpha \leq 0.05$.

Table 3: Pre-Test Writing Score of the Students in the Controlled and Experimental Groups

Group	N	Mean	Std. Deviation	t	Sig.(2-tailed)
Controlled Group	47	35.94	4.03	-.45	.651
Experimental Group	47	36.30	3.68		

However, after the treatments, it was found that students who were taught with Contrastive Analysis Grammar Instruction showed a greater improvement in writing ability to those who were taught in a traditional way. As shown in Table 4, the t-value was calculated to be -8.10, which is statistically significant at $\alpha \leq 0.05$.

Table 4: Post-Test Writing Score of the Students in the Controlled and Experimental Groups

Group	N	Mean	Std. Deviation	t	Sig.(2-tailed)
Controlled Group	47	37.21	3.62	-8.10	.000
Experimental Group	47	43.85	4.30		

As illustrated in the examples below, before this student received Contrastive Analysis Grammar Instruction, she had made a wide range of errors, such as improper uses of the definite article, pronouns, verbs, prepositions, and auxiliary verbs; missing verbs, prepositions, conjunctions, articles, verb "be", and the subject; redundant verbs, pronouns, prepositions, and verb "be"; noun forms; word choices; word positions; subject and verb agreement; misspelling; and meaningless noun phrases. However, after the treatment, she made a few errors which are only misspelling and an improper use of a reflexive pronoun.

A. Experimental Group

Before

My Favorite Pet

"My favorite pet is *cat* [noun form]. Cat has *the* [improper use of 'the'] round big eyes, which [lack of a verb] no harm. When I make eye contact [lack of a preposition] *he* [improper use of pronoun], I feel *to be* [redundant verb] happy. Cat *is* [incorrect verb use] characteristics *such as* [word choice], he *cry* [subject and verb agreement] [lack of a conjunction] [lack of an article] baby *me* [redundant pronoun]. When I come *to* [improper use of preposition] home, he is playful. When I play with *he* [improper use of pronoun], *sometime* [misspelling], he [lack of a verb "be"] shy when [lack of a subject] find a newcomer or stay *in* [redundant preposition] there *many people* [meaningless noun phrase]. *The most* [word position] of cat

favorite food is fish, especially *is* [redundant verb "be"] mackerel. But occasionally he *like* [subject and verb agreement] to have a dessert such as jelly and [lack of an article] kind of Thai sweet meat. He *stay* [subject and verb agreement] with me when I *don't be* [improper use of auxiliary verb] pleasure and lonely, therefore, I love him."

After

The Animal in My Heart

"The animal in my heart is a cat. They have four characteristics that make me like them. First of all, cats are very cute. Their eyes are big and beautiful. They are playful and friendly. Second, cats are small. You can take a cat to go shopping and for a walk. Also, you can carry a cat everywhere you want. Third, cats are economical. You don't have to buy a cat because you can take cats from temples or your neighbor's house. In addition, you don't have to feed a cat because they can go hunting by *itself* [improper use of reflexive pronoun]. When they are hungry, they eat rats in your house. Finally, cats are clean. They make your house clean from rats. In *conclution* [misspelling], if you feed a cat, you will feel happy and relaxed because they are cute, playful, friendly, small, economical and clean."

Compared to the examples above, the students in the controlled group did not show much improvement in their English writing ability. As shown below, after the treatment, this student still made a lot of errors, such as improper uses of verbs, nouns, and pronouns; redundant articles, verbs, and prepositions; missing auxiliary verbs, possessive adjectives, and verb "be"; subject and verb agreement; verb forms; incomplete sentences; punctuation marks; and verb tenses.

B. Controlled Group

Before

My Favorite Pet

"My favorite pet is *parrot* [noun form] that is a good pet. It has many *advantage* [plural noun]; for instance, it can talk that [improper use of relative pronoun] the parrot is so talkative, [comma splice] it's so beautiful and it can fly. *Sometime* [misspelling] if you're bored, you can play with *parrot* [noun form]*that* [improper use of relative pronoun] you feel good because the parrot can talk with *yours* [improper use of possessive pronoun] so you are so happy but the parrot has many *disadvantage* [plural nouns]; for example, it's so noisy when the parrot *talk* [subject and verb agreement] or *sing* [subject and verb agreement] a song *in* [improper use of preposition] the [improper use of "the"] night because you can't sleep that [improper use of relative pronoun] you are enduring. In conclusion, if you feed the parrot, you [lack of a modal] have knowledge about *advantage* [plural noun] and *disadvantage* [plural noun]."

After

The Animal in My Heart

"*Dog are* [subject and verb agreement] cutest things. *I have ever seen.* [incomplete sentence] Because they have something in them. First, all of *dog* [plural noun] *have* [improper use of verb] *a* [redundant article] honest and royal. When *they* [improper use of pronoun] owner *have* [redundant verb] leave *from* [redundant preposition] home and come back. *Dog* [plural noun] *like* [subject and verb agreement] to *jumping* [verb form] and *rolling* [verb form] on the ground because they [lack of an auxiliary verb] happy to see [lack of a possessive adjective] owner. Second thing [lack of a verb "be"] dogs have soft fur. I like to grope them and dogs have a good smell. Last thing [lack of a verb "be"] when I look in their eyes [lack of a comma] I *saw*

[verb tense] love from them that [lack of a verb "be"] why dogs *had become* [verb tense] the animal in my heart."

According to the statistical evidence and the students' pieces of writing, it is clearly seen that the experimental group showed a greater improvement on writing ability than the controlled group. Apart from this, the students in the experimental group had fewer writing errors compared to those in the controlled group.

Decreases in Writing Errors

As shown in Table 5, it is apparent that the number of both grammatical and mechanical errors of the students in the experimental group has decreased from the total of 554 errors in the pre-test to 209 errors in the post-test. For the grammatical features taught in the treatment (i.e. subject-verb agreement, active and passive voices, word order in noun phrases and sentences, and verb tenses) in particular, the students have shown their great improvement on those features. Before the treatment, subject and verb agreement had been the most commonly found error after articles (a, an, the); however, the number was reduced from 34 errors to only 8 errors. Similar to the other features, the erroneous uses of active and passive voices became less from 11 errors to 8 errors, word order in the noun phrase from 26 errors to 14 errors, verb tenses from 70 errors to 30 errors and sentence structure from 57 errors to 17 errors.

Table 5: Types of Writing Errors Found in the Pre-Test and Post-Test Written Pieces of Students in the Experimental Group

Types of Writing Error	Pre-Test		Post-Test	
	Frequency (Number of Word Errors)	**Percentage (%) Compared to Other Writing Errors**	**Frequency (Number of Word Errors)**	**Percentage (%) Compared to Other Writing Errors**
Nouns	12	2.2	4	1.9
Transitive and Intransitive Verbs	15	2.7	3	1.4
There + be	32	5.8	6	2.9
Auxiliary Verbs	9	1.6	4	1.9
Modal Verbs	7	1.3	8	3.8
Adverbs of Frequency	7	1.3	7	3.3
Possessive Determiners	13	2.3	4	1.9
Prepositions	18	3.2	8	3.8
Articles	70	13	14	6.7
Pronouns	14	2.5	14	6.7
Relative Pronouns	7	1.3	4	1.9
Infinitives	3	0.5	3	1.4
Gerunds	0	0	2	1.0
Conjunctions	6	1.1	2	1.0
Any/Some	6	1.1	2	1.0
Word Order in the Sentence	9	1.6	4	1.9
Word Order in the Noun Phrase	26	4.7	14	6.7

Types of Writing Error	Pre-Test		Post-Test	
	Frequency (Number of Word Errors)	Percentage (%) Compared to Other Writing Errors	Frequency (Number of Word Errors)	Percentage (%) Compared to Other Writing Errors
Word Meanings	13	2.3	8	3.8
Subject-Verb Agreement	34	6.1	8	3.8
Incomplete Sentences	19	3.4	5	2.4
Run-on Sentences	29	5.2	8	3.8
Present Simple	11	2.0	7	3.3
Present Perfect	8	1.4	3	1.4
Past Simple	13	2.3	6	2.9
Future Simple	8	1.4	4	1.9
Active and Passive Voices	11	2.0	8	3.8
Unparalleled Verb Tenses	30	5.4	10	4.8
Redundant Verbs	28	5.1	0	0
Redundant Nouns	19	3.4	7	3.3
Redundant Pronouns	16	2.9	0	0
Redundant Articles	11	2.0	0	0
Redundant Prepositions	8	1.4	4	1.9
Spelling	24	4.3	14	6.7
Capitalization	10	1.8	5	2.4
Punctuation	8	1.4	9	4.3
Total	554	100	209	100

Compared to the experimental group, the students in the controlled group made less progress in English grammar and mechanics. As shown in Table 6, even though the number of grammatical and mechanical errors has decreased from the total of 590 errors in the pre-test to 358 errors in the post-test, they made more errors in some grammatical points in the post-test. They made 19 run-on sentences and 4 present perfect errors in the pre-test, but it was 20 and 5 errors in the post-test, respectively.

Table 6: Types of Writing Errors Found in the Pre-Test and Post-Test Written Pieces of Students in the Controlled Group

Types of Writing Error	Pre-Test		Post-Test	
	Frequency (Number of Word Errors)	Percentage (%) Compared to other Writing Errors	Frequency (Number of Word Errors)	Percentage (%) Compared to other Writing Errors
Nouns	19	3.2	14	3.9
Transitive and Intransitive Verbs	14	2.4	13	3.6
There + be	35	5.9	26	7.3
Auxiliary Verbs	5	0.8	4	1.1
Modal Verbs	3	0.5	8	2.2
Adverbs of Frequency	13	2.2	5	1.4
Possessive Determiners	10	1.7	3	0.8
Prepositions	25	4.2	19	5.3
Articles	72	12	40	11

Types of Writing Error	Pre-Test		Post-Test	
	Frequency (Number of Word Errors)	Percentage (%) Compared to other Writing Errors	Frequency (Number of Word Errors)	Percentage (%) Compared to other Writing Errors
Pronouns	10	1.7	12	3.4
Relative Pronouns	7	1.2	4	1.1
Infinitives	5	0.8	0	0
Gerunds	1	0.2	0	0
Conjunctions	14	2.4	11	3.1
Any/Some	2	0.3	2	0.6
Word Order in the Sentence	11	1.9	7	2.0
Word Order in the Noun Phrase	23	3.9	19	5.3
Word Meanings	18	3.1	7	2.0
Subject-Verb Agreement	30	5.1	18	5.0
Incomplete Sentences	14	2.4	12	3.4
Run-on Sentences	19	3.2	20	5.6
Present Simple	19	3.2	16	4.5
Present Perfect	4	0.7	5	1.4
Past Simple	15	2.5	12	3.4
Future Simple	4	0.7	3	0.8
Active Voices and Passive	11	1.9	9	2.5
Unparalleled Verb Tenses	38	6.4	16	4.5
Redundant Verbs	25	4.2	17	4.7
Redundant Nouns	11	1.9	0	0
Redundant Pronouns	12	2.0	0	0
Redundant Articles	25	4.2	13	3.6
Redundant Prepositions	5	0.8	0	0
Spelling	50	8.5	19	5.3
Capitalization	15	2.5	0	0
Punctuation	6	1.0	4	1.1
Total	590	100	358	100

The information above shows that the English writing progress of the students who received Contrastive Analysis Grammar Instruction was better than the controlled group and that raises the question what exactly in the instruction helped them enhance their writing ability. The following part gives the answer revealing the students' attitudes toward the instruction and what made Contrastive Analysis Grammar Instruction beneficial for Thai EFL students.

Interview Results

Most of the interviewees claimed that the English and Thai languages had many different features that made English language learning difficult for Thai students. Many students use Thai grammar when they write English sentences. They usually have trouble using verb tenses, subject and verb agreement, passive and active voices, and inflectional suffixes. Because verb tenses are contained in the English language but they are not in Thai, when Thai students write English sentences, they usually use base forms with all tenses. Even with the present simple tense, Thai students usually forget to change the verb form to correspond with the subject of the sentence. There is no suffix in Thai, but

there are too many suffixes in English. A verb in English shows the time from its form. The "-ed" suffix or irregular form is added to the verb to show the time of the event. Readers know whether the event happens in the present, past or future from the verb form. They even know who is the subject or the object of the sentence in active and passive voices. In addition, the English suffixes appear when making adjectives from verbs, verbs from nouns, and adverbs from adjectives. For example, the noun and adjective "summary" can change its function when "-ize" is added to make it a verb "summarize" or "ily" to make it an adverb "summarily". However, the mentioned language features are not contained in the Thai language. Thai students who are not familiar with the differences usually have trouble writing English. They omit the "-ed" ending when they talk about actions in the past. They generalize the "-ed" ending with all verbs very often. They cannot identify parts of speech from word forms. Another point needs considering; in terms of the language structures, although both English and Thai are SVO, word orders in the noun phrase of the two languages are totally different.

"If Thai students do not understand about the differences of English and Thai languages miscommunication can occur. For example, most of Thai students trend to use Thai structure when they write in English..."

"Thai students taking an English course are not familiar with the English structure…"

"The English language has somewhat different structure from the Thai language, and that can make difficulty for Thai students. In other words, they will use English roughly because of such a difference…"

"Some grammatical features in English such as subject and verb agreement, verb tenses and passive voice do not exist in the Thai language…"

"Tenses help students know about the time and know which event comes first or which comes after…"

In Thai, only adverbs are added to show he past form while in English, the form of a verb can be changed…"

"In English, the suffix "ed" is added to the verb to show the past form…"

"In English verbs can change their forms, like adding "es", "s", "ed" or irregular verbs. However, such a method doesn't exist in Thai…"

"Subject and verb agreement helps students understand the role of the subject…"

"Thai students are not familiar with using English suffixes, such as [summary], [summarize] and [summarily]..."

"Passive voice makes students know who is the agent or who is the object of the sentence..."

"The arrangement of words in English phrases is different from those in Thai. We say [two big boxes] in English, but it is [Krong Yai Songbai] or [boxes big two] in Thai…"

Ironically, with Contrastive Analysis Grammar Instruction, trouble with writing English sentences was reduced. The students claimed that when they learned about which grammatical features were similar to or different from their native language, they could write English sentences more correctly and suitably. For those which were similar to Thai, like the sentence structure (Subject-Verb-Object or SVO), they used it in their pieces of writing without much attention. For the different features, they had to be more careful and pay more attention to when they wrote. That's to say, they knew which ones

require more time to comprehend before use. Such a method helped them memorize the English grammatical features faster. Apart from this, when they knew the English grammar well, the number of writing mistakes was reduced. So far as that is concerned, it was always kept in mind that they had to change verb forms for each tense in English. They knew that the verb needed to agree with the subject of the sentence and words in the English noun phrase were placed differently from Thai.

"Students will not use Thai structure when they write in English because they learn about the differences of the two languages."

"Contrastive analysis makes students learn which grammatical features are similar with Thai language."

"Students can learn how to use the English language suitably and fast. We can also analyse the event from grammatical features. Our mistakes can be reduced if we knew English grammar well. It is always struck in our mind that we had to add "ed" or change verb forms. We know the verb must follow the subject. We know that the English noun phrase is positioned differently from the Thai noun phrase. When we know the similarities and differences between the two languages, it is struck in our mind and helps us realize how to write English more accurately."

According to the interview results, English and Thai have many different features and some aspects of English are not contained in Thai. Such a large number of differences have brought about a lot of difficulties for Thai students to learn English. However, with the aid of Contrastive Analysis Grammar Instruction, those difficulties were gotten rid of. Not only did the instruction raise the students' awareness of language differences when they used some unique features in English but also save their time when they learned those similar ones.

Discussion and Conclusion

The present study reveals that Contrastive Analysis Grammar Instruction improved writing ability of Thai EFL pre-intermediate students. As is proven by the pre-test and post-test writing scores, the students who received Contrastive Analysis Grammar Instruction showed a greater improvement in writing ability than those who were taught with the traditional grammar instruction at the significant level in statistics. The instruction also had positive effects in terms of the enhancement of the students' writing ability and attitudes toward learning English. Apparently, their pieces of writing were more comprehensible and accurate after the treatment. They made fewer writing mistakes and were much likely to write English in a proper way following Standard English Grammar. Considering the interview responses, it was found that the students had positive attitudes toward Contrastive Analysis Grammar Instruction. They claimed that this teaching method helped reduce writing mistakes because positive transfer, noticing similarities between the two languages, helped them learn faster while negative transfer, learning from different features, was a reminder that they had to pay more attention to different English grammatical features when they wrote.

As mentioned above, Contrastive Analysis Grammar Instruction is one of the effective language teaching methods that should be promoted. Not only did it reduce writing mistakes but the process also helped students think carefully about how to use the English

grammatical features which were different or nonexistent in Thai. Such an answer yields support to beliefs of scholars and evidence of previous studies that have mentioned about the advantages of Contrastive Analysis in language teaching. According to Lado (1957), with Contrastive Analysis, similarities and differences between first and second languages are emphasized and that helps the teacher predict language learning difficulties of students; such an approach leads to learning success. Schuster (1997) asserts that similar features between the two languages can make language learning easier while those different or nonexistent features like in the case of English and Asian languages can cause much learning trouble. Especially for advanced learners, König and Gast (2009) claim that students can learn better from a direct comparison of their mother tongue with the target language. Ghabanchi and Vosooghi (2006) found that the students taught through Contrastive Analysis had significantly higher scores on both recognition and the production tasks than those who studied English grammar from comprehensible input only.

As is proven in this article, it is undeniably true that Contrastive Analysis Grammar Instruction has been an effective grammar teaching. With its usefulness in the language classroom, students can easily recognize different grammar features making them learn the target language better.

Implications on Institutional Research and Education

Future Research

The present study has proven that Contrastive Analysis Grammar Instruction helped Thai EFL students improve their writing ability. For future research, researchers who are interested in this teaching method can use it to teach other language skills, such as speaking, listening, reading or even integrated skills. As discussed in this study, Contrastive Analysis Grammar Instruction worked well with undergraduate students; besides, future researchers can make it more challenging with other levels of students, namely high school, secondary school, or primary school.

Classroom Management

Before using Contrastive Analysis Instruction in the classroom, the teacher needs to prepare a good learning environment. When students want some clarification, they must feel free to ask questions from the teacher. Besides, the teacher should survey students' needs to find out what grammatical features they consider important for their learning because students who have different levels of language proficiency need different inputs. For example, primary school students can learn to use verb "be" quickly, but for some other complex features like gerunds or infinitives, they can be inappropriate for their level of proficiency. Most importantly, when the teacher compares the two languages, the examples given to students must start with the less complex structure to the more complex one. Such a step can help students notice the similarities and differences between the languages more easily.

Acknowledgements

I thank Ms. Pim Gengsumrong and all the participating students for their contribution to this research.

References

Abbas, A. (1995). Contrastive analysis: Is it a living fossil? *IRAL: International Review of Applied Linguistics in Language Teaching, 33*, 195-215.

Bennui, P. (2008). Study of L1 interference in the writing of Thai EFL students. *Malaysian Journal of ELT Research, 4*, 72-102.

Bootchuy, T. (2008). *An analysis of errors in academic English writing by a group of first-year Thai graduates majoring in English.* Master of Arts Thesis, Kasetsart University.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing.* New York: Cambridge University Press.

Corder, P. (1967). The significance of learners' errors. *IRAL: International Review of Applied Linguistics in Language Teaching, 5*, 161-170.

Fisiak, J. (1981). Some introductory notes concerning contrastive linguistics. In J. Fisiak (Ed.), *Contrastive linguistics and the language teacher* (pp. 1-13). Oxford: Pergamon Press.

Ghabanchi, Z., & Vosooghi, M. (2006). The role of explicit contrastive instruction in learning
difficult L2 grammatical forms: A cross-linguistic approach to language awareness. *The Reading Matrix, 6*, 144-153.

Hellinger, M., & Ammon, U. (1996). *Contrastive sociolinguistics.* Berlin: Mouton de Gruyter.

House, J. (1996). Contrastive discourse analysis and misunderstanding: The case of German and English. In M. Hellinger & U. Ammon (Eds.), *Contrastive sociolinguistics* (pp. 345-361). Berlin: Mouton.

James, C. (1980). *Contrastive analysis.* London: Longman.

James, C. (2005). Contrastive analysis and the language learner. In D. J. Allerton, C. Tschichold, & J. Wieser (Eds.), *Linguistics, language teaching and language learning* (pp. 1-20). Basel: Schwabe.

Jawasi, N. (2006). *Contrastive linguistics.* Medan: Program Pasca Sarjana UNIMED.

Khamput, P. (2004). *First language interference on second language writing of Mattayomsuksa 5 students.* Master of Arts: An Independent Study Report in English, Khon Kaen University.

Khaoural, P. (2002). *A study of error analysis in English composition: Case study of English major students of Rajabhat Institute Nakhon Pathom.* Master of Arts Thesis in Language and Culture for Communication, Mahidol University.

Klein, W. (1986). *Second language acquisition.* Cambridge: Cambridge University Press.

König, E., & Gast, V. (2009). *Understanding English-German contrasts* (2nd Ed.). Berlin: Erich Schmidt Verlag.

Kortmann, B. (1996). Kontrastive linguistik und fremdsprachunterricht. In W. Börner & K. Vogel (Eds.), *Kontrast und aquivalenz. beiträge zu sprachvergleich und übresetzung* (pp. 136-167). Tübingen: Narr.

Lado, R. (1957). *Linguistics across cultures.* Ann Arbor: Michigan University Press.

Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research.* London: Longman.

Likitrattanaporn, W. (2002). Ponkarnwikrohkhobokprongtangwaiyakornphasa Angkristornaewtangkaekhaipanhakarnriankarnsorn. (An analysis of grammatical errors towards approaches to solutions of instructional problems). *Journal of Humanities Parithat, Srinakharinwirot University, 24*, 80-90.

Mair, C. (2005). Recent advances in contrastive linguistics and language typology: The spin-off for language teachers. In D. J. Allerton, C. Tschichold, & J. Wieser (Eds.), *Linguistics, language teaching and language learning* (pp. 21-39). Basel: Schwabe.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology.* Mahwah, NJ: Lawrence Erlbaum Associates.

Pongpairoj, N. (2002).Thai university undergraduates' errors in English writing. *Journal of Languages and Linguistics, 20,* 66-99.

Nemser, W. (1971). Approximative systems of foreign language learners. *IRAL: International Review of Applied Linguistics in Language Teaching, 9,* 115-123.

Noochoochai, P. (1978). *Temporal aspect in Thai and English: A contrastive analysis.* Doctor of Philosophy Thesis in the School of Education, Health, Nursing and Arts, New York University.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning.* Cambridge: Cambridge University Press.

Oller, J., & Ziahosseiny, S. (1970). The contrastive analysis hypothesis and spelling errors.
Language Learning, 20, 183-189.

Pengpanich, A. (2002). *Error analysis of English usage and use* (5th Ed.). Bangkok: Ramkhamhaeng University Press.

Pongpairoj, N. (2002).Thai university undergraduates' errors in English writing. *Journal of Languages and Linguistics, 20,* 66-99.
Ratanakul, S. (2000). Verb-tense errors in English writing made by Thai students:
A comparative analysis. *Journal of Language and Culture, 19*, 77-85.

Ridwan, T. A. (1998). *Dasar â€“ dasar linguistik kontrastif.* Medan: USU Press.

Rungnaphawet, A. (2005). A contrastive analysis of adverbs in English and Thai, and its implication on language teaching. *PASAA, 36*, 89-114.

Schuster, K. (1997). *Second language acquisition: Lado and the contrastive analysis.* Retrieved from http://www.coas.uncc.edu/linguistics/courses/4263/4263lado.htm

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*, 209-231.

Sheen, R. (1996). The advantage of exploiting contrastive analysis in teaching and learning a foreign language. *IRAL: International Review of Applied Linguistics in Language Teaching, 34*, 183-197.

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly, 4*, 123-130.

Wierzbicka, A. (1985). Different cultures, different languages, different speech acts. *Journal of Pragmatics, 9.2*, 145-178.

Wierzbicka, A. (1992). *Cross-cultural pragmatics: The semantics of human interaction.* Berlin: de Gruyter.

Appendix

Scoring Rubric: Writing Ability

Description	Maximum score	Earned score
1. Format (3 points)		
- There is a title.	1	
- The title is centered.	1	
- There is only one indentation.	1	
Total		
2. Organization (10 points)		
- The paragraph begins with a topic sentence that has both a topic and a controlling idea.	3	
- The paragraph contains several supporting sentences.	4	
- The paragraph ends with an appropriate concluding sentence.	3	
Total		
3. Content (10 points)		
- The paragraph fits the assignment and is relevant to the topic.	3	
- The paragraph is interesting to read.	3	
- The paragraph shows coherence and unity.	4	
Total		
4. Punctuation and Mechanics (10 points)		
- Spelling	3	
- Capitalization	3	
- Commas, periods and other punctuation marks	4	
Total		
5. Grammar and Sentence Structure (10 points)		
- Correct sentence structure with no fragments	3	
- Various types of sentences (Simple, compound and complex sentences)	3	
- Overall grammar (The grammatical points which has been already taught in class)	4	
Total		
6. Overall Impression (7 points)		
GRAND TOTAL	50	

Opinion Page

WHAT LEARNING SYSTEM DO WE NEED WITH THE EVER RAPID AND RATCHETED CHANGES

Nirwan Idrus
nirwan.idrus@gmail.com

H G Wells (1866-1946), a British writer once wrote:
"Adapt or perish, now as ever, is nature's inexorable imperative."

The idea of systems that are adaptive, that is able to change as requirements vary, is only logical in this day and age where technology has and continues to *mess-up* the complacent folks. The idea of contentment has indeed become anachronistic, perhaps even during H G Wells' life for otherwise he would not have written the quote above. We have heard that the only constant is in fact change.

A decade or so ago, manufacturing technology researchers and practitioners sought to find the seemingly elusive manufacturing system that could become the panacea for the potential manufacturing problems coming on stream.

The only thing wrong with that was the assumption that there was only going to be one change to face and this would occur infrequently and far between. Henry Ford survived despite his famous utterance that you can only choose Ford Model T and that it is black, because there was no competition and he was the one who came up with the various creative ideas such as the assembly line.

Frederick W. Taylor circa 1850 came up with Scientific Management which atomized each task in a job and thus either applicable to or in fact became a preamble to Henry Ford's assembly line. A chicken and egg story one may say.

While it was undoubted that both Taylor and Ford had opened people's eyes on the need for efficiency and effectiveness, like anything else it was not enough. Customers' expectations are naturally elevated each time they experienced a good thing like quality, cheaper product or services and so on.

The recognition of this asymptotic phenomenon had got researchers back to their drawing boards and looked for something else perhaps more permanent and able to automatically adjust to the ever increasing demands. Hence the creation of concepts such as Artificial Intelligence, Integrated (and Computer Integrated) Manufacturing, Continuous Quality Improvement, Kaizen, Jidoka and even Adaptive Manufacturing.

In fact manufacturing is not the only industry that is affected by the groaning and pervasive development caused primarily by technology advancement. In a tragic twist, education is equally affected. Tragic, because education is the prime mover of advancement and not affected by something that it supposedly caused to happen.

The tragedy with education generally and in higher education specifically, is not only that it has lost, if it ever has, the role of education as society would expect it, but for the bulk of laggards and shirkers when it comes to improving or modernizing aspects of education.

This is of course not saying that nothing has been done to improve education. Indeed there has been a lot done around the world on improving education. Literature would show umpteen education reforms in a country. Have they made any difference to their respective education? To be sure, it would be difficult to say no. Developing countries would celebrate the results of their education reforms because access to education in the country has increased or that new schools or universities were established irrespective of their quality. The sophistication of financial manipulation of foreign aids in developing countries were so high that reports of poor building and facilities quality in those newly built amenities could and did trigger another source of personal income for some. And that is only the quality of the buildings. What about the quality of the syllabi, teachers, lecturers and professors? They may have been mentioned in the education reform documents, but what can you do when there is not anybody around to assume those roles let alone being of quality. This does not mean that there are not the people in those developing countries who do not aspire to be quality academics and academic administrators.

Then come technology advancement, innovation, parents' and students' expectations which kept increasing and creating new demands, as well as a realization that there is so much to education than just a teacher teaching and directing students on what to do. Generation challenges, either brought about by technology advancement or not, become a challenge to education, to teachers, to syllabi, to equipment and everything else education. But that is only one.

Another formidable challenge of course is that everyone is an expert in education. This is the bad news for education particularly in developing countries inhabited by defensive

people. Instead of admitting their shortfalls and trying to resolve these challenges, they would expend unnecessary time to prove that suggestions for improvement by those who know are inferior to what they are doing, be it proven useless 150 years ago. Imagine if they had spent that time improving themselves.

Quite fortunately, those persistent educationalists though not recognized nor appreciated, continue to find the missing pieces to make education relevant to individual's, society's and national needs.

Tim Hudson (2015) recently published a White Paper on Adaptive Learning. Why? Because of the challenges above. He calls it Adaptive because it is not only meeting the idiosyncrasies of today's but also of the outcomes of meeting them. He recognized that in any classroom, there is one teacher and many learners. The learners in turn are made up of a multitude of backgrounds, intelligence, general knowledge, commitments to learning and academic abilities. Thus, some learners need least help in class while others need a close handholding. Importantly, there is no learning model that would accommodate these differences in students. This is eloquently expressed by Hudson as:

> "• As class sizes increase, teachers have less time to observe students as they work individually.
>
> • A single group of 15 to 30 students may have a range of abilities and preparedness — from English language learners and students working below grade level to accelerated learners ready to zoom ahead.
>
> • There is no single model of instruction that works for all students, and the number of instructional strategies any single instructor can know and use is limited.
>
> • The promise of data-driven decision making to individualize instruction is hampered by the lag between assessment, reporting, analysis, and action."

De'javu says the quality people who have had to contend with a multitude of quality definitions. Just like education, quality is something that everybody claims to know well. While that is a story in itself and too big to be included in a short article such as this, it suffices to say that quality, like education, must be treated in its totality, from awareness to building people's confidence to admit that they don't know everything to 'there is a better way'... and to involve everyone, yes everyone from top management, to administrators, to academics to students, future students, their supports, the government, the industry, the creative as well as the innovative. This very list is not exhaustive of course and perhaps has missed a number of important protagonists as it is.

Intelligent Adaptive systems are rapidly becoming the order of the day. The word itself conjures an unending series of activities with one aim and one aim only, namely giving full satisfaction to the clients, customers and stakeholders whoever they may be, including but perhaps especially students. Education adds uniqueness in that our customers or clients are also the *raw material* that we are charged to manipulate in order

to produce useful, informed, creative, innovative graduates some of whom for all we know may in fact dedicate themselves and their work to the betterment of education in the country. Given this, education would seem suitable to adopt Adaptive learning system.

In order to implement Adaptive Learning in education, Hudson listed the following considerations:

> "• **Degree of adaptation.** Does the technology offer a few paths with student-directed pacing or does it offer millions of learning paths with fine-grained adaptation?
>
> • **Rigor of the curriculum.** Does the technology only focus on practice or does it offer a rich environment for developing conceptual understanding, computational fluency, and problem solving?
>
> • **Level of engagement.** Does the technology appeal to a narrow range of students or does it empower all students with choices and personalized challenges that help them take ownership of their learning experience?
>
> • **Quality of reports.** Does the technology only report time on task and completion or does it give teachers real-time analysis of comprehension and class-wide data to help with grouping and differentiation?"

Further information on the technology is available in Hudson's paper cited below.
This paper looks a little deeper at the implementation of such technology in developing countries, like Malaysia. I believe that anything can be accomplished when the *people* are willing to not only implement the technology for the sake of implementation but are committed to improving themselves as well.

A major obstacle seems to be the old and anachronistic belief held by many that students don't know anything and the teacher/lecturer knows everything. Of course it was never true, simply because humans are not God, so there is no way that a human being, be him/her is a teacher or not, knows everything. Those who subscribe to this idea are obviously uninformed and have no idea on how to inform themselves. The big question is...Do you want your kids to be taught by such persons? Given the proliferation of information and the increasing number of data sources, it is obvious that that sort of lecturers is the worst thing you can have in education.

Another aspect that appears to need fundamental change is the teaching/learning style in many of these countries. Perhaps it is not their fault, but perhaps it is, but *rote learning* appears to be the dominant teaching/learning style despite the fact that it was declared ineffective in the late 1800's. The question again is: Do you want to send your kids to study at a university that practices a 150-year old and ineffective teaching/learning style? The associated assessment method may give your kid an "A" or several "A"s but you know that s/he cannot think for themselves, nor have initiatives and are simply misaligned with the sorts of skills and cognitive requirements of our ever changing world.

Last but not least, though there are lots more in between, is compartmentalization. It is no secret that one department in a university does not want to know about the happenings in another department, nor do they share information that only needs to be acquired once and yet we, the students and other stakeholders are being asked more than a dozen times by the same organization for the same information, and in some cases, have to be in printed form that would then need to be (slow) mailed at a time when the world breathes instant everything.

So, it is very important to have a thorough preparation before adopting a new technology, learning style or anything else that one may one day accuse it to be useless simply because you cannot implement it properly in the first place. This thorough preparation should concentrate on:

- Own capability compared to the capabilities required
- Own capability to change
- Own willingness to change, the first one being able to accept that rather than you know everything, to you know little and others know more than you
- Own willingness therefore to unlearn all the previous beliefs and learn all the new ones that will help improve education overall
- Restructuring the organization by de-compartmentalizing first followed by the accompanying management skills such as communicating, doing, trying, risking and the like

These bullet points have been the actions that many who wanted to move ahead did not do and hence whatever technology new and old that is introduced to the classrooms and even management of the university or faculty never did get off the ground. It is true of course that a lot of these shortfalls is the result of poor leaderships, but good leadership will not transpire unless individuals either collectively or individually do the bullet points above in the first place.

Hudson's (2015) Adaptive Learning technology proves to be more than just addressing learning style in classrooms as the bullet points above proved that adaptive approach like Quality is required in all aspects of an endeavour.

Reference:

Hudson T (2015) *Help all students excel in Math with Adaptive Learning Technology*, White Paper WP0026 DreamBox Learning Inc.

Bibliography:

Campus Technology (2012) *Technology Transforming Education: 4 Real-World Models of Success,* White Paper Rosetta Stone Education http://www.RosettaStone.com/highereducation Retrieved 13/1/2013

Chen W Y and Idrus N (2008) *Towards a synergized multi-level effective education: The Secondary School-University Alignment Workshop – an initiative for effective and informed teaching and learning* Procs SEAAIR Annual Conference 2008, November 2008, STIE Perbanas, Surabaya, Indonesia

Choi A S (2014) *What the best education systems are doing right* ideas.ted.com (Explore ideas worth spreading, every weekday) 09 April 2014

Idrus N (2008) *Quality Learning in Higher Education: A book about Practical Quality Improvement* ISBN 978-974-615-295-2, Assumption University Press, Bangkok, Thailand

Idrus N (2008) *A preliminary study on reading habits of students and staff in Asian HEIs: Towards a balanced global K-Capacity blueprint* .Procs SEAAIR Annual Conference 2008, November 2008, STIE Perbanas, Surabaya, Indonesia

Idrus N (2009) *Comment: Quo Vadis Higher Education?* JIRSEA vol. 7 no. 1, 2009, 71-73

Idrus N (2011) *Challenges facing Higher Education as a service organization in a developing country.* International Journal of Services, Economics and Management 2011 - Vol. 3, No.1 pp. 104 - 121

Idrus N (2014) *South East Asia Higher Education at the Cross Roads* Procs SEAAIR International Conference, Legazpi, The Philippines, http://www.seaairweb.info/Conference Retrieved 13/10/2014

Idrus N, Ng P K, Jee K S (2014) *Sustaining Quality in Higher Education in Southeast Asia through Understanding Generational Changes* J App Sc 14 (16) 1819-1827
Ministry of Education (2012) *Malaysian National Graduate Employability Blueprint 2012-2017,* Government of Malaysia