

JAPANESE UNIVERSITY ENGLISH LANGUAGE TEACHERS' SELF-EFFICACY BELIEFS: A MIXED-METHODS EXPLORATION

A Dissertation
Submitted
to the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education

By
Max Praver
May, 2014

Examining Committee Members:

David Beglar, Advisory Chair, Teaching and Learning
Eton Churchill, External Member, Kanagawa University
Kay Irie, External Member, Tokai University
Lance Burrows, External Member, Kinki University
James Elwood, External Member, Meiji University

©
by
Max Praver
2014
All Rights Reserved

ABSTRACT

Japanese University English Language Teachers' Self-Efficacy Beliefs:

A Mixed-Methods Exploration

Max Prayer

Doctor of Education

Temple University, 2014

Doctoral Advisory Committee Chair: Dr. David Beglar

This study is an investigation of Japanese university English language teachers' self-efficacy beliefs. Research has established that teachers' self-efficacy has considerable influence on a wide variety of teaching practices. However, in the English as a Foreign Language domain, and more specifically at the university level in Japan, self-efficacy beliefs have hardly ever been examined.

The purpose of this study was to investigate teachers' self-efficacy beliefs based on the teachers' native language, teaching experience, contract and tenured status, and gender. Furthermore, the sources of these beliefs, how they are strengthened, and how they are challenged were also explored.

In order to provide answers to these questions, the Japanese University Language Teachers' Efficacy Beliefs Scale (JULTEBS), a new instrument measuring language teacher self-efficacy was validated using the Rasch rating-scale model as well as a confirmatory factor analysis. A triangulation strategy mixed-method design was employed in which the collection and analysis of data from the quantitative survey was

completed in addition to the collection and analysis of data from qualitative open-ended interviews. A profile analysis, a special application of a MANOVA, was conducted to check the hypotheses for parallelism, levelness, and flatness of the self-efficacy scores among the various groups of respondents. The four self-efficacy variables that were measured were Efficacy in Student Engagement, Efficacy in Instructional Strategies, Efficacy in Classroom Management, and Efficacy in Dealing with Superiors. Semi-structured interviews were also employed to help determine what potentially strengthens and weakens the self-efficacy beliefs of English language teachers.

The results showed that native English language teachers perceived themselves to be more efficacious than Japanese English teachers across all four self-efficacy variables. Additionally, more experienced teachers exhibited higher self-efficacy beliefs than less experienced teachers. Tenured teachers and limited-term contract teachers showed similar levels of self-efficacy on all variables except for Efficacy in Dealing with Superiors, where tenured teachers rated themselves higher than contract teachers. Furthermore, male and female teachers showed no statistically significant differences across all four self-efficacy variables. Finally, four themes (Autonomy, Colleagues, Money, and Students) emerged as qualities that could support teachers' self-efficacy, whereas three themes (Administration, Students, and Limited-term Contracts) surfaced as qualities that could weaken teachers' self-efficacy.

The findings of this study not only highlight the importance of teacher self-efficacy, but also provide valuable insights into the beliefs of English language teachers, as well as the current state of affairs for these teachers at Japanese universities.

ACKNOWLEDGMENTS

Without the support of numerous individuals, the completion of this dissertation would not have been possible.

First of all, I would like to thank Dr. David Beglar for supporting me throughout the entire dissertation writing process. He provided continual and helpful feedback from day one. I cannot thank him enough for his relentless generosity, encouragement, and guidance. He was truly amazing!

I would also like to thank Dr. Eton Churchill, Dr. Kay Irie, Dr. Lance Burrows, and Dr. Jim Elwood for serving as members on my committee at a very busy time of the year. Their rigorous scrutiny of my work, the challenging questions they posed during the defense, and their insightful comments provided me the opportunity to significantly improve my dissertation. Dr. Jim Elwood deserves extra special thanks for his many hours helping me understand the intricacies of not only the Rasch model but also EQS and structural equation modeling. Without his kindness, well-timed comic relief, and baseball commentary, I fear I would have completed this thesis slightly more scarred.

It was also a great pleasure to share this gargantuan undertaking with my cohort mates both in Osaka and Tokyo. Their friendly encouragement through what seemed to be endless hours of study and writing was instrumental in pushing me forward.

Last but not least, I would like to thank both of my families here in Japan and also in the United States. Chiemi, Emma, Mom, Dad, Sam, and Meredith: I could not have finished my dissertation without your understanding and support. Thank you!

TABLE OF CONTENTS

	Page
ABSTRACT	iv
ACKNOWLEDGMENTS	vi
LIST OF TABLES	xiii
LIST OF FIGURES	xv
CHAPTER	
1. INTRODUCTION	1
Background of the Issue.....	1
Statement of the Problem.....	5
Purposes and Significance of the Study.....	7
Theoretical Perspective: Social Cognitive Theory	8
Audience for the Study	9
Researchers.....	9
Administrators and Instructors	10
Pre-Service Faculty	10
Delimitations.....	11
Definitions of Terms.....	12
Organization of the Study.....	12
2. REVIEW OF THE LITERATURE.....	14
Social Cognitive Theory	14
Triadic Reciprocal Determinism	16
Human Agency.....	17
Self-Efficacy	18

Self-Efficacy and Outcome Expectancy.....	19
Self-Concept, Self-Esteem, and Locus of Control	20
Self-Efficacy and Motivation	23
Self-Efficacy and Motivated Teachers	23
Self-Efficacy and Teacher Anxiety.....	26
Self-Efficacy and Teacher Burnout	27
Self-Efficacy and Gender	29
Sources of Efficacy Beliefs	32
Teachers’ Self-Efficacy and its Measurement	34
Measurement of Teacher Efficacy Based on Rotter’s Theory	35
Rotter’s Theory of Locus of Control	35
The Rand Studies	36
Teacher Locus of Control	37
Webb Efficacy Scale.....	38
Measurement of Teacher Efficacy Based on Bandura’s Theory.....	38
Gibson and Dembo’s Teacher Efficacy Scale	38
Conceptual Confusions	39
Bandura’s Teacher Self-Efficacy Scale	42
An Integrated Model	44
Teachers’ Sense of Efficacy Scale (TSES)	46
Subject-Matter and Context-Specific Scales	50
Science Teaching	50
Special Education	51
Classroom Management	52

English as a Foreign Language Teachers' Self-Efficacy	53
TESOL and Self-Efficacy	53
Gaps in the Literature.....	56
Purposes of the Study.....	59
Research Questions.....	60
3. METHOD.....	61
Mixed-Methods Research Design.....	61
Participants.....	63
Questionnaire.....	63
Interviews	66
Ethics	67
Instrumentation	67
Procedures.....	71
Interviews.....	71
Analyses.....	73
Quantitative Analyses.....	74
Rasch Analysis.....	74
Confirmatory Factor Analysis	78
Profile Analysis.....	81
Qualitative Analyses.....	84
Summary.....	86
4. INSTRUMENT VALIDATION	87
Rasch Analysis of Item Fit and PCA of Item Residuals.....	87
Pilot Study Analysis.....	88

Efficacy in Dealing with Superiors	89
Efficacy in Student Engagement	89
Efficacy in Classroom Management	90
Efficacy in Instructional Strategies	91
Summary of Pilot Study Rasch Analysis	91
Data Screening	92
Instrument Validation: Rasch Analysis	93
Construct 1: Efficacy in Dealing with Superiors.....	93
Construct 2: Efficacy in Student Engagement	98
Construct 3: Efficacy in Classroom Management.....	102
Construct 4: Efficacy in Instructional Strategies.....	106
Summary	110
Instrument Validation: Confirmatory Factor Analysis	111
Confirmatory Factor Analysis Procedures	111
Confirmatory Factor Analysis Results	113
Respecification of the Model	113
Instrument Validation Summary	117
5. QUANTITATIVE RESULTS	123
Research Question 1: Self-Efficacy and Native Language.....	122
Research Question 2: Self-Efficacy and Teaching Experience	125
Research Question 3: Self-Efficacy and Contract Status.....	129
Research Question 4: Self-Efficacy and Gender	132
Summary	134
6. QUALITATIVE RESULTS	136

Interviews.....	136
Participant Characteristics	136
Patterns and Themes.....	139
Research Question 5: Potential Self-Efficacy Supports.....	140
Autonomy	140
Teaching Style	141
Curriculum and Class Design	141
Decision Making.....	143
Colleagues	144
Knowledge Sharing.....	144
Cooperation.....	146
Money.....	148
Research.....	150
Salary	151
Students	152
Research Question 6: Potential Self-Efficacy Challenges	156
Administration.....	156
Lack of Communication	157
Poor Management	158
Students	161
Contract Teachers.....	165
Summary	169
7. DISCUSSION	170
Social Cognitive Theory and Reciprocal Determinism	170

Profile Analysis.....	173
Research Question 1	174
Research Question 2	181
Research Question 3	186
Research Question 4	192
Contributions and Implications.....	196
8. CONCLUSION	202
Summary of the Study	202
Limitations of the Study.....	204
Suggestions for Future Research	206
Final Conclusions.....	209
REFERENCES	211
APPENDICES	
A. ONLINE CONSENT FORM	225
B. BACKGROUND QUESTIONNAIRE.....	226
C. JAPANESE UNIVERSITY LANGUAGE TEACHER EFFICACY BELIEFS SCALE.....	229
D. THE 42 SELF-EFFICACY ITEMS USED IN THE PILOT STUDY	232
E. THE 32 SELF-EFFICACY ITEMS EXTRACTED FROM THE PRELIMINARY ANALYSIS	234
F. THE FINAL 28 POST-INSTRUMENT VALIDATION SELF-EFFICACY ITEMS USED IN THE MAIN ANALYSIS.....	235
G. INTERVIEW CONSENT FORM	236
H. GUIDING INTERVIEW QUESTIONS	237

LIST OF TABLES

Table	Page
1. Teacher Self-Efficacy Measures, Scale Types, and Example Items	44
2. The Participants' Native Language and Gender.....	65
3. The Participants' Teaching Experience and Gender	65
4. The Participants' Contract Versus Tenure Status and Gender	65
5. The Participants' Gender.....	65
6. Japanese University Language Teacher Efficacy Beliefs Scale (Initial Version).....	70
7. Rating Scale Instrument Quality Criteria	78
8. A Summary of the Four Factors Identified by the Pilot Study Questionnaire	92
9. Category Structure Functioning for Dealing with Superiors	94
10. Rasch Item Statistics for Efficacy in Dealing with Superiors (Pre-Item Deletion).....	94
11. Rasch Item Statistics for Efficacy in Dealing with Superiors (Post-Item Deletion).....	95
12. Category Structure Functioning for Student Engagement	99
13. Rasch Item Statistics for Efficacy in Student Engagement.....	99
14. Category Structure Functioning for Classroom Management	103
15. Rasch Item Statistics for Efficacy in Classroom Management.....	103
16. Category Structure Functioning for Instructional Strategies	107
17. Rasch Item Statistics for Efficacy in Instructional Strategies	107
18. A Summary of the Four Factors Identified by the Questionnaire	112
19. A Summary of the Final Number of Respondents	120
20. A Summary of the Final Items from the Questionnaire Used in the Main Analysis.....	120

21. The Final Japanese University Language Teacher Efficacy Beliefs Scale	121
22. Descriptive Statistics for Native and Non-Native English Teachers Across the Four Self-Efficacy Factors	123
23. Profile Analysis for the Four Self-Efficacy Factors by Native Speaker Group	123
24. Descriptive Statistics for Teaching Experience For the Self-Efficacy Factors	126
25. Profile Analysis for the Four Self-Efficacy Factors by Experience	127
26. Descriptive Statistics for Tenured and Limited-Term Contract English Teachers Across the Four Self-Efficacy Factors	129
27. Profile Analysis for the Four Self-Efficacy Factors by Employment Status	130
28. Descriptive Statistics for Male and Female English Teachers Across the Four Self-Efficacy Factors	132
29. Profile Analysis for the Four Self-Efficacy Factors by Gender	133
30. Summary of the Profile Analyses Results	135
31. Teacher Bio-Information	137
32. Main Categories from the Interview Data	140

LIST OF FIGURES

Figure	Page
1. Bandura's model of triadic reciprocal determinism	9
2. Teachers' sense of efficacy integrated model.....	44
3. Triangulation strategy design.....	62
4. Exponential non-discriminative snowball sampling.....	63
5. Timeline of the creation and administration of the JULTEBS	69
6. Wright map for Dealing with Superiors	97
7. Wright map for Student Engagement	101
8. Wright map for Classroom Management.....	105
9. Wright map for Instructional Strategies.....	109
10. Hypothesized four-factor university English language teacher self-efficacy model	112
11. Standardized coefficients for the four-factor self-efficacy model.	114
12. Final version of the four-factor self-efficacy model.....	118
13. Profiles of the four self-efficacy factors for native English teachers and Japanese English teachers.....	125
14. Profiles of the four self-efficacy factors for teachers' years of experience	128
15. Profiles of the four self-efficacy factors for the teachers' employment status	131
16. Profiles of the four self-efficacy factors for the teachers' gender	134

CHAPTER 1

INTRODUCTION

The Background of the Issue

Japanese universities have arrived at an interesting point in time. Forecasted for some time now, this era is not only intriguing but also disconcerting in several ways. First, the demographic trend of declining birthrates has created a situation where the number of college-aged young people has decreased by so much that the country has reached a period of full college and university admissions (Hani, 2001; White, Eguchi, Kawanaka, & Henneberry, 2005). In other words, any young Japanese person who wishes to pursue a post-secondary degree will have a 'seat' at a university somewhere in the country. Although the idea of everyone being able to attend university sounds quite encouraging, in fact, there are two very real and disastrous implications to this situation. The most obvious repercussion is the lack of tuition, in other words, income that will not be available to those universities that are unable to fill their admission quotas. The top-tiered universities, due to their popularity and reputations will almost always be able to fill their admission target (Fitzpatrick, 2010). However, it is unavoidable that falling enrollment caused by declining birthrates will eventually all but cripple the lower level universities. As enrollment decreases and funds dry up, the harsh economic effects will be felt at all levels and in all departments of the affected institutions (Sawa, 2004). The second and perhaps not as obvious repercussion is the predicament of how to meet admission quotas and face the increasing lack of competition for admissions to many institutions. By eliminating competition and offering a place to all applicants, a general

decline in the overall academic abilities of incoming college students is inevitable.

Institutions quite simply will have to lower their standards and admit students that would not have been accepted even a decade ago (Hani, 2001).

Furthermore, despite the implications stated above, and ostensibly undeterred by falling scholastic aptitude and quickly vanishing wealth, for at least the past 20 years there has been a growing push to ‘internationalize’ the moribund Japanese education system (Fitzpatrick, 2010). This push to internationalize has been operationalized in a variety of ways, such as the implementation of the Japanese Exchange and Teaching (JET) program in the late 1980’s (McConnell, 2000) and the attempt to introduce English language classes to elementary schools (McCurry, 2011) during the last 10 years. One of the most recent attempts to address the issue of internationalization, the ‘Global 30’ project, introduced by the Japanese Ministry of Education in 2009 (McNeil, 2010), has been implemented at the university level. Under this project, by the end of 2013, more than 130 undergraduate and graduate courses are to be launched and conducted completely in English at 13 universities acting as Japan’s ‘global education hubs.’ The project is designed to not only draw foreign students and teachers to Japanese universities, no doubt trying to counter their diminishing student populations, but also to better prepare Japanese students to study abroad. Despite having praiseworthy goals, the Global 30 project is off to a shaky start with enormous government budget cuts already eliminating 30% of allocated funds to the 13 universities (Aoki, 2010).

The aforementioned issues present a dilemma for the Japanese educational system. With lofty goals of internationalizing Japan, conflicting with the reality of a declining student population, a lack of government funds, and the decreasing academic

ability of its young people, Japanese universities have reached a perplexing crossroads. Stuck in the middle of this conflict, a part of the population often overlooked, are the teachers working at post-secondary institutions in Japan. More specifically, they are the English language faculty who are essential to the internationalization process. It is these teachers who are expected to respond to the demands of the government and raise the level of English of the now less-able students, while simultaneously working in less economically stable conditions. Despite the unique circumstances facing the Japanese educational system and the apparent crucial role that these teachers are to play in the government's over-arching educational plan, little research has been conducted into the beliefs and attitudes of these language teachers.

One type of belief that has gained popularity in educational psychology research over the past four decades and that is informative when examining teachers' attitudes and capabilities is that of self-efficacy. Research has established that teachers' self-efficacy beliefs have a beneficial impact on a wide variety of features that are important to both teaching and learning (Henson, 2002; Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998; Woolfolk Hoy, Davis, & Pape, 2006). Studies have revealed that it is positively associated with students in terms of their motivation (Midgley, Feldlaufer, & Eccles, 1989), achievement (Ross, 1992, 1995; Tschannen-Moran, et al., 1998), and sense of efficacy (Anderson, Greene, & Loewen, 1988). Self-efficacy is also positively associated with teachers in terms of their classroom management (Shim, 2001; Woolfolk & Hoy, 1990; Woolfolk, Rosoff, & Hoy, 1990), job satisfaction (Caprara, Barbaranelli, Borgogni, & Steca, 2003), enthusiasm for teaching (Allinder, 1994), planning and organization (Allinder, 1994; Milner, 2001), and

persistence in the face of difficulties (Milner, 2002; Milner & Woolfolk Hoy, 2003; Ross, 1998). This research makes self-efficacy an ideal construct to use when investigating the English teaching situation in Japan. With the overall declining ability of college-aged students, now more than ever, teachers with higher levels of confidence and self-efficacy will be necessary in order to deal with and impart knowledge on a generally less able student body.

Teachers' self-efficacy beliefs, also commonly labeled *teacher efficacy*, *teachers' self-efficacy*, or *teachers' sense of efficacy*, are defined as "the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context" (Tschannen-Moran & Woolfolk Hoy, 2001, p. 223). Accordingly, teachers' sense of efficacy can be understood as self-perceived beliefs about their ability to successfully carry out their teaching tasks in their specific teaching contexts.

Concurrent with the escalation of research into general teacher self-efficacy, researchers have also explored teachers' sense of efficacy in specific subject areas (Bleicher, 2004; Chacón, 2002, 2005; Coladarci & Breton, 1997). Among other subjects, for example, researchers in the field of science education have investigated teachers' sense of efficacy in teaching the subject matter in different contexts (Bleicher, 2004; Riggs & Enochs, 1990; Roberts & Henson, 2000). However, in the field of Teaching English to Speakers of Other Languages (TESOL) and more specifically at the university level, inquiry into teachers' sense of efficacy is exceedingly limited (Chacón, 2002; Shim, 2001). Given its documented influence on teaching practices and student learning (Henson, 2002; Tschannen-Moran, et al., 1998; Tschannen-Moran & Woolfolk Hoy,

2001; Woolfolk Hoy, et al., 2006), further pursuit of this line of investigation is warranted.

Statement of the Problem

This study is designed to address two problems. Despite the well-documented importance of fostering teacher self-efficacy and in light of its predictive benefits to instructors and students, the examination of teacher self-efficacy at universities in Japan has never been addressed. Bandura (1997) cautioned against assessing self-efficacy by using measures that do not consider domain specific dimensions, and in fact, trying to measure the self-efficacy of Japanese university teachers at this point in time would prove futile, as no appropriate instrument exists. Teachers, for example, who judge themselves highly efficacious teaching English grammar in a junior high school classroom with 40 students, might be much less self-assured of their efficacy to teach a 15-person university seminar on English debate, and vice versa. In other words, the first problem is that there is no currently available measure of teacher self-efficacy that encapsulates not only the work, but also the work environment at Japanese universities. The development of such a questionnaire is challenging, given that self-efficacy is not uniform across different subjects or domains of instructional functioning (Bandura, 1997), and a self-efficacy measure designed for one type of teacher in a specific context is not necessarily appropriate for another. Furthermore, adding to the aforementioned limitations of the currently available self-efficacy scales is the often-overlooked fact that teacher efficacy measures have developed a controversial history regarding construct validity and measurement integrity. The construct validity of the primary instruments

purporting to measure teacher efficacy have been severely criticized (Denzine, Cooney, & McKenzie, 2005; Guskey & Passaro, 1994; Henson, 2002). Accordingly, teacher efficacy is presently at the proverbial ‘crossroads’; it is ready to either progress or to fade away as a seemingly good idea that in the end had little to offer. As Tschannen-Moran et al. (1998, p. 202) noted:

This appealing idea, that teachers’ beliefs about their own capacities as teachers somehow matter, enjoyed a celebrated childhood, producing compelling findings in almost every study, but it has also struggled through the difficult, if inevitable, identity crisis of adolescence...teacher efficacy [now] stands on the verge of maturity.

As such, it is clear that theoretically as well as psychometrically sound instruments are not only needed, but necessary if the field of teacher self-efficacy is to move forward.

The second problem is that the self-efficacy of English instructors at Japanese universities has never been explored in depth. An examination of currently available teacher self-efficacy scales highlights their inability to assess the population of interest in this study. Current assessments are designed to measure the self-efficacy beliefs of teachers not only outside of postsecondary instruction, but outside of language instruction, as well. For example, the majority of current teacher self-efficacy scales have been developed, calibrated, and validated using only general elementary and secondary level teachers (Ashton, Olejnik, Crocker, & McAuliffe, 1982; Dellinger, Bobbett, Oliver, & Ellett, 2008; Gibson & Dembo, 1984; Rose & Medway, 1981; Tschannen-Moran & Woolfolk Hoy, 2001). In light of these considerations and as no such assessment tool has been developed, not only is the introduction of an instrument measuring the self-efficacy beliefs of English instructors at Japanese universities warranted, but the concurrent measurement of the self-efficacy beliefs of these instructors is also justified.

Purposes and Significance of the Study

There are two purposes to this study. The first purpose is to introduce and validate an instrument appropriate for investigating Japanese university English teachers' self-efficacy beliefs. This is significant because no instrument currently exists that fully captures the intricacies of the jobs of university English teachers. By developing and validating such an instrument, it will fill this gap in the field of TESOL and add another valuable tool to the self-efficacy research repertoire.

The second purpose is to explore self-efficacy beliefs of university English teachers in Japan. To achieve this objective, I employ a mixed-methods design, more specifically a triangulation strategy (Creswell, 2009). As such, the study consists of two phases. The collection and analysis of the questionnaire (quantitative data) in the first phase is followed by the collection and analysis of the interviews (qualitative data) in the second phase. In this study weight is given to the questionnaire data. In other words, I first use the newly created instrument to compare and contrast various groups of university English teachers' self-efficacy levels. By this, I mean to look at the (a) the teachers' native language, (b) years of teaching experience, (c) tenured vs. contract status, and, (d) gender, and determine if the groups have parallel scores on the self-efficacy measure. Second, in order to provide an important supplement to the quantitative data, I conduct qualitative interviews to explore teacher supports (or lack of) that influence the teachers' self-efficacy beliefs. Exploring self-efficacy beliefs of university English teachers in Japan is significant in several ways and can make contributions to expanding the current state of knowledge in the TESOL field. First, the study addresses the need to inquire into university teachers' beliefs in their capability for teaching English by

adopting the notion of teacher self-efficacy. Although many educational researchers have documented the powerful impacts of teachers' sense of efficacy on various student and teacher aspects in teaching and learning (Henson, 2002; Tschannen-Moran, et al., 1998), only a few researchers (Chacón, 2002; Shim, 2001) have studied it in the TESOL field. In addition, no researchers have investigated university teachers' sense of efficacy in an EFL setting such as Japan. Given the distinctiveness of the environment at Japanese universities and the importance of teacher development in that context, it is valuable to investigate the levels of teachers' self-efficacy to provide at the very least, a baseline picture of the current status of university English education in Japan from the teachers' perspective. The present study is also significant because of the methodology employed. There has been an emergent need for qualitative inquiry in addition to quantitative inquiry in the teacher efficacy research (Henson, 2002; Labone, 2004). As Henson (2002) stated: "...because an efficacy judgment is a result of an individual's filtering of internal and external factors, the context surrounding a person's judgment is very relevant to the study of teacher efficacy, but this is captured poorly by Likert-type questionnaires" (p. 147). The design of the current study takes this point into consideration, as the interviews contribute to contextualizing the survey findings.

Theoretical Perspective: Social Cognitive Theory

In this study, I explore the notion of self-efficacy through Bandura's (1986; 1997) social cognitive theory. Social cognitive theory, in which self-efficacy plays a central role, assumes that people are capable of the intentional pursuit of courses of action (human agency), and that such agency operates in a process called triadic reciprocal causation. In

other words, social cognitive theory defines human behavior as an interaction among personal factors, behavior, and the environment. According to this theory, an individual's actions are uniquely determined by each of these three factors. Each part of this trinity (see Figure 1) mutually impacts the other two parts. Bandura also suggests this trinity determines what people come to believe about themselves, and affects the choices they make and actions they take. Not surprisingly, as the central construct in social cognitive theory, Bandura (1997) explains that self-efficacy beliefs can influence the courses of action people choose to pursue, how much effort they put forth in given endeavors, and how long they persevere in the face of obstacles and failures.

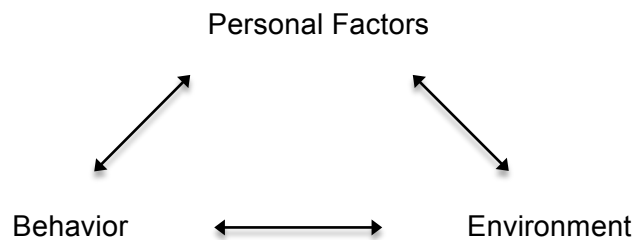


Figure 1. Bandura's conceptual model of Triadic Reciprocal Determinism.

The Audience for the Study

Researchers

Because self-efficacy is thought to be domain specific, the exploration of self-efficacy beliefs in the Japanese university context makes a valuable contribution to the broader field of self-efficacy research. Furthermore, it introduces a new theoretically and psychometrically robust scale, the Japanese University Language Teachers' Efficacy

Beliefs Scale (JULTEBS), to the list of available self-efficacy inventories, and will be of interest to those researchers involved in scale development.

Administrators and Instructors

Nearly every institution of higher education in Japan today employs numerous English teachers and offers a variety of English course options. A standardized scale that evaluates the self-efficacy beliefs of English instructors can potentially provide the administrators the means to identify where improvement in for example, curriculum and administration, and faculty support is needed. If the proper infrastructure is in place, professional development programs, faculty mentoring, instructional workshops, and more ideal working conditions can be developed to improve English teaching at Japanese universities. By diagnosing potential challenges and difficulties, efforts toward solutions and better educational practices can be pursued.

Pre-Service Faculty

The audience for this study is not only limited to current faculty, but also extends to future faculty enrolled in graduate programs. In identifying the self-efficacy beliefs of Japanese university English teachers, administrators and policy makers will gain a greater understanding as to what components need to be included in graduate education and training programs. This understanding can provide a means through which postsecondary institutions and graduate programs can better prepare future faculty for university English teaching roles.

Delimitations

All of the data were gathered from English language teachers working at public and private two and four-year accredited postsecondary institutions in Japan during 2011. Stratified purposeful and snowball sampling strategies were utilized. Therefore, while the population to which generalizations can be made is to English language teachers working at universities in Japan, generalizations about the self-efficacy beliefs of teachers of younger students (e.g., junior or senior high school teachers) or English language teachers working outside of Japan must be made with caution. Furthermore, as the sampling strategies employed in this study did not create a random sample of the population, a potentially skewed data set is conceivable and generalizations must be made with caution.

A final delimitation to this study involves the questionnaire items measuring the Dealing with Superiors construct and the ability of part-time teachers to accurately assess these items. Many part-time university teachers in Japan have little contact with their superiors and therefore might have difficulty conceiving of the situations described in the instrument. Additionally, part-time faculty who teach at several institutions might find the questionnaire items difficult to respond to due to differential treatment and teaching environments they experience at their respective schools. The results to this part of the study should also be interpreted cautiously.

Definition of Terms

The following three terms are frequently used in this study:

1. Self-efficacy: “beliefs in one’s capabilities to organize and execute a course of action required to produce a given attainment” (Bandura, 1997, p. 3).
2. Teacher Efficacy: In the present study, teacher efficacy refers to “the teacher’s belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context” (Tschannen-Moran et al., 1998, p. 233). The term *teacher efficacy* is interchangeably used with *teachers’ sense of efficacy* and *teachers’ self-efficacy*.
3. Japanese University Language Teachers’ Efficacy Beliefs Scale (JULTEBS): Teacher efficacy in teaching English is operationalized as the individual mean scores on the four dimensions extracted from the 28-item Japanese University Language Teachers’ Efficacy Beliefs Scale (JULTEBS) which was modified from the Teacher Sense of Efficacy Scale (Tschannen-Moran & Woolfolk, 2001) to fit the specific context of teaching English at Japanese universities.

The Organization of This Study

There are eight chapters in this study. Chapter 2 is a review of the literature in which social cognitive theory, self-efficacy, and teacher self-efficacy beliefs are reviewed, the gaps identified through the literature review are described, and the research questions are presented. The participants, development of the instrumentation, procedures, and the analyses used in this study are described in Chapter 3, Methodology. A report of the preliminary analyses in which the questionnaire was tested and analyzed using the Rasch

rating-scale model and a Confirmatory Factor Analysis to verify the validity and reliability is presented in Chapter 4, Instrument Validation. After the preliminary analyses and instrument validation, quantitative and qualitative results are reported in Chapters 5 and 6, respectively. Key findings, theoretical implications, and pedagogical implications are discussed in Chapter 7, Discussion. In the final chapter, Conclusion, the limitations of the study, suggestions for future research, and final conclusions are presented.

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of this chapter is to provide a comprehensive review of previous research about social cognitive theory, self-efficacy, and teachers' self-efficacy, while also presenting the gaps in those fields, and posing research questions. The chapter is divided into three major sections: Social Cognitive Theory, Self-Efficacy, and Teachers' Self-Efficacy. The gaps in the literature, purposes of the study, and the research questions are provided at the end of the chapter.

Social Cognitive Theory

The theoretical groundwork of self-efficacy is located in social cognitive theory, developed in 1977 by Albert Bandura. Social cognitive theory (Bandura, 1977, 1986, 1997) explains human behavior in terms of a triadic reciprocal interaction of personal, behavioral, and environmental influences. In other words, people function as contributors to their own motivation, beliefs, and behavior within a network of reciprocally interacting factors. Bandura labeled this theory as "cognitive" to stress the important influence of cognition in people's capability to encode information, self-regulate, and perform behaviors.

Bandura's (1997) social cognitive theory differs from behaviorist theories conceiving human change as the product of environmental or external stimuli. Unlike the behaviorist perspectives, in social cognitive theory, human change cannot be reduced to the result of external stimuli, because human thoughts also influence behaviors through

introspection. Similarly, Pajares (2002) also pointed out that Bandura's social cognitive theory differs from the theories overemphasizing the influences of biological factors on human change and adaptation, because they fail to adequately consider social and contextual influences. Thus, social cognitive theory equally emphasizes human agency and environmental influences in conceptualizing human change and adaptation, rejecting a duality between human agency and social structure (Bandura, 1997).

Bandura's concept of self-efficacy, the focus of this paper, was developed as part of a larger theory, the social learning theory (Ashford & LeCroy, 2010), which has progressed into the social cognitive theory (Levin, Culkin, & Perrotto, 2001). As alluded to above, Bandura presented social cognitive theory in response to his dissatisfaction with the principles of behaviorism and psychoanalysis. In these two theories, the role of cognition in motivation and the role of the situation are largely ignored (Bandura, 1977). Bandura (1977) summed this up nicely when he explained,

Unidirectional environmental determinism is carried to its extreme in the more radical forms of behaviorism...but humanists and existentialists, who stress the human capacity for conscious judgment and intentional action, contend that individuals determine what they become by their own free choices. Most psychologists find conceptions of human behavior in terms of unidirectional personal determinism as unsatisfying as those espousing unidirectional environmental determinism. To contend that mind creates reality fails to acknowledge that environmental influences partly determine what people attend to, perceive, and think. (pp. 344-345)

Nevid (2009) explains that social cognitive theory illustrates the fact that individuals do not simply respond to environmental influences, but rather they actively seek and interpret information. Individuals "function as contributors to their own motivation, behavior, and development within a network of reciprocally interacting influences" (Bandura, 1999, p. 169). Although social cognitive theory covers many topics

such as moral judgment and physiological arousal, most research is primarily focused on self-efficacy, or the beliefs regarding one's capabilities of successfully completing tasks or goals (Locke & Latham, 2002). According to Bandura (2005), social cognitive theory takes on an agentic perspective to change, development, and adaptation. Bandura describes an agent as someone who intentionally influences one's functioning and life circumstances. He states: "In this view, people are self organizing, proactive, self-regulating, and self reflecting. They are contributors to their life circumstances not just products of them" (Bandura, 2005, p. 1).

The key assumptions of social cognitive theory, including reciprocal determinism, human agency and its capabilities, are explained in the following section.

Triadic Reciprocal Determinism

Social cognitive theory assumes that human behavior, the environment, and personal factors mutually interact and serve as determinants of each other. However, this principle of triadic reciprocal determinism (see Figure 1) does not imply that these factors affect each other simultaneously or equally. The strength of influence depends on activities, individuals, and circumstances (Bandura, 1986).

A bi-directional interaction occurs between behavior and personal factors in the sense that because people's beliefs, expectations, and goals shape their behavior, the consequences of their behavior also influence their personal characteristics. The personal factors-environment interaction of reciprocal determinism is also a two-way interaction. Not only are people's expectations, beliefs, and cognitive competencies developed and altered by their environment, but they also influence their environment. Finally, the

mutual interaction between behavior and environment suggests that people are both producers and products of their environment (Bandura, 1986, 1997).

Human Agency

A principal piece of reciprocal determinism is the concept of human agency. Agency refers to “acts done intentionally” (Bandura, 1997, p. 3). Social cognitive theory assumes that people have power to influence and make changes in their actions. Bandura (1997) emphasized that efficacy beliefs are the most influential characteristics of human agency. He suggested, “unless people believe they can produce desired effects by their actions, they have little incentive to act. Efficacy is, therefore, the foundation of agency” (pp. 2-3).

Social cognitive theory assumes that people have a number of fundamental competencies that distinguish them as a human (Bandura, 1986, 1989):

1. People have *symbolizing* capabilities that enable them to provide their lives with form, meaning, and persistence. For Bandura, “symbols serve as the vehicle of thought” (Bandura, 1989, p. 9). Through symbols, such as mental images or words, they can develop new courses of action by testing possible solutions hypothetically. This capability can also allow for creation of internal models that guide future actions and communication with others.
2. People have *forethought* capability, that is, they are able to set goals and anticipate the likely consequences of actions. Through these expectations, people choose actions likely to produce desired outcomes rather than detrimental ones.
3. People can learn *vicariously* by observing other’s actions and the consequences of their

actions. This capability allows people to develop new behavior while avoiding the trial-and-error process of actually performing it.

4. People are capable of *self-regulation* that allows them to have personal control over their own motivations and actions. They develop personal standards, evaluate their performance against these standards continuously, and thus motivate themselves to work harder and change behavior in their succeeding actions.

5. People are *self-reflective*, that is, they analyze and evaluate their experiences and their own thought processes. By engaging in self-evaluation, they change their behavior and thinking accordingly.

With five assumed competencies that distinguish people as humans and countless factors that potentially influence human functioning, human agency is central to the potential for self-directed changes in behavior. Within the socio-cognitive framework, the most vital regulatory mechanism for behavior is an individual's judgments of their capabilities to "organize and execute courses of action required to produce given attainments" (Bandura, 1997, p. 3), or one's self-efficacy beliefs.

Self-Efficacy

Self-efficacy beliefs are at the very core of social cognitive theory and since the publication of Bandura's article "Self-efficacy: Toward unifying theory of behavior change" in 1977, this notion has spawned a growing body of literature not only in education, but also in many other fields, such as medicine and business administration.

Bandura (2006) has conceived self-efficacy beliefs as the most central mechanism of human agency in social cognitive theory. In relation to this, he stated:

Among the mechanisms of human agency, none is more central or pervasive than belief of personal efficacy. This core belief is the foundation of human agency. Unless people believe they can produce desired effects by their actions, they have little incentive to act, or to persevere in the face of difficulties. Whatever other factors serve as guides and motivators, they are rooted in the core belief that one has the power to effect changes by one's actions. (p. 170)

Self-efficacy beliefs refer to personal judgments about one's capabilities to perform actions at designated levels. Efficacy beliefs are concerned not with the number of skills you have, but with what you believe you can do with what you have under a variety of circumstances (Bandura, 1997).

Self-Efficacy and Outcome Expectancy

Self-efficacy theory also proposes that outcome expectations form a second construct related to motivational behavior and affect. Outcome expectations are judgments or beliefs regarding the likelihood between a person's behavior and the anticipated outcome. This notion of contingency between response and outcome is similar to Rotter's (1966) construct of locus of control regarding the likelihood between behavior and reinforcement. In addition, in terms of the anticipation of success, it is similar to expectancy for success from expectancy-value theories. As Bandura (1986) puts it, "The belief that one can high jump six feet is an efficacy judgment; the anticipated social recognition, applause, trophies, and self-satisfactions for such a performance constitute the outcome expectations" (p. 391). In the academic domain, students would have efficacy judgments of their capabilities, skills, and knowledge to master school-related tasks, but also have outcome expectations about what grades they might receive on the tasks. Although efficacy beliefs and outcome expectations are

usually positively correlated, it is possible for a student to have a relatively high efficacy belief for a task, but low outcome expectations. For example, a Japanese college student in an English writing class might have relatively high efficacy beliefs about personal capability to master the material, but low outcome expectations about grades on exams due to the very high competition among other English students and the grading curve instituted by faculty to weed out the weaker students.

Although Bandura proposed both of these motivational constructs (self-efficacy and outcome expectations), the theory and subsequent research focus on the role of self-efficacy beliefs. Bandura (1986) suggests that outcome expectations are heavily dependent on efficacy judgments: “If you control for how well people judge they can perform, you account for much of the variance in the kinds of outcomes they expect” (p. 393). Bandura notes that outcomes are connected to actions; how one behaves largely determines the actual outcome and, in the same way, beliefs about outcome expectations are dependent on self-efficacy judgments. He gives the example that drivers who are not confident in their ability to negotiate a winding mountain road (low efficacy) will conjure up images of wreckage and injuries (one type of outcome expectation), whereas those confident in their ability will anticipate the grand views from the mountains.

Self-Concept, Self-Esteem, and Locus of Control

Bandura (1997) distinguished self-efficacy from other similar constructs such as self-concept, self-esteem, and locus of control. He pointed out that although they are self-referential, self-efficacy is different from all other self-constructs in that it involves judgments of capabilities specific to a particular task.

Self-concept is a global construct that contains many perceptions about the self, only one of which is self-efficacy. Self-concept refers to a composite view of oneself that is formed through direct experience and evaluations adopted from significant others (Bandura, 1997); thus, it is mostly concerned with global self-images. Unlike self-concept, self-efficacy beliefs vary according to the domain of activities, the levels of difficulty, and the specific context. For example, self-concept might be too global in nature to distinguish someone who has low self-efficacy in teaching English grammar to large classes of junior high school students. Because of the specificity property of self-efficacy, Bandura (1997) proposed that it is a better predictor of behavior than self-concept.

Self-esteem refers to perceptions of self-worth and does not include judgments of capabilities like self-efficacy. There is no preset relationship between beliefs about one's capabilities and whether one likes or dislikes oneself. Therefore, people's judgment of their capacity to perform a certain activity as low does not necessarily involve a loss of self-esteem. For example, one's belief about their capability to teach English grammar to junior high school students does not impact their self-esteem as a tennis coach, unless of course the person invests their sense of self-worth in their tennis coaching abilities to how well they teach English grammar.

Although locus of control and self-efficacy are viewed as similar constructs, they correspond to entirely different phenomena (Bandura, 1997). Bandura explained that locus of control is an outcome expectancy that can be defined as a person's estimate that a given behavior will lead to certain outcomes. High locus of control, however, does not necessarily indicate a sense of empowerment and wellbeing. For example, professional

baseball players might believe that a high salary is completely dependent on their performance (high locus of control), but because they believe they lack the skills to produce those superior performances (low self-efficacy), feel hopeless.

In sum, self-efficacy is specific to a domain, the level of difficulty within the domain, and the context. These aspects differentiate self-efficacy beliefs from other self-referent constructs, such as self-concept, self-esteem, and locus of control.

Fittingly, the prominence of self-efficacy beliefs can clearly be seen in Pajares' (2002) statement:

How people behave can often be better predicted by the beliefs they hold about their capabilities than by what they are actually capable of accomplishing, for these self-efficacy perceptions help determine what individuals do with the knowledge and skills they have. (p. 4)

While a disparity between one's self-judgments and truth is rather typical, self-judgments often occur first when one engages in a course of action (Pajares, 2002). Therefore, Pajares explains that people's attainments are generally better predicted by their self-efficacy beliefs than by their previous attainments, knowledge, or skills. Similarly, Bandura (1997) elucidated that people's level of motivation, affective states, and actions are more based on what they believe than what is objectively true. This displays the importance of self-efficacy beliefs in human functioning.

In terms of how self-efficacy beliefs influence human functioning, Bandura (2006) has contended that they influence people's goals and aspirations, how well they motivate themselves, and their perseverance in the face of difficulties and adversity. Bandura also believes that self-efficacy beliefs shape people's outcome expectations and determine how opportunities and impediments are viewed. For example, a person with

high efficacy might be more resilient in the face of adverse situations while a person with low efficacy might give up more easily.

Self-Efficacy and Motivation

Aside from being connected to other self-related concepts as explained above, self-efficacy and motivation also share much in common. For example, it is not difficult to imagine that teachers who possess high levels of self-efficacy do more for their students and to be generally more content with their teaching situations. Nor is it difficult to conceive that teachers with low levels of self-efficacy might suffer from teacher anxiety or teacher burnout. Self-efficacy has been found to be a powerful predictor of various constructs and can act as a strong antecedent to an individual's motivation, both positively and negatively. This section aims to situate self-efficacy amongst three common affective phenomena, and to clarify their respective relationships.

Self-efficacy and motivated teachers.

Williams and Burden (1997) have called motivation a multi-faceted construct and defined it as “a state of cognitive and emotional arousal, which leads to a conscious decision to act, and which gives rise to a period of sustained intellectual and/or physical effort in order to attain a previously set goal (or goals)” (p. 120). Much like Bandura's (1997) concept of self-efficacy, it appears that teacher motivation might play a significant role in affecting student motivation, and be an important component in the classroom. Unfortunately, also like self-efficacy research, it appears that a majority of the research in the field of TESOL has focused on student motivation, and there has been much less

research on teacher motivation. Dörnyei (1998) explained the scarcity of language teacher motivation research,

Teacher motivation has been a largely uncharted area in the L2 [second language] field...As far as I am aware, no L2 study has explicitly linked the level of teacher motivation with that of students, and the topic of teacher motivation has also received little attention in general educational psychology. This is all the more surprising since the teacher's level of enthusiasm and commitment is one of the most important factors that affect the learners' motivation to learn. (p. 130).

Many early theories viewed motivation as a reaction to factors outside human control.

This led to cognitive approaches to motivation, which saw choice as the main factor.

Williams and Burden (1997) broadened the cognitive perspective into what they call a social constructivist view, which emphasized the individual uniqueness of motivation.

This view is similar to Bandura's (1977) social cognitive theory in the sense that people's motivation is influenced by social and contextual factors and they make choices to act based on personal reasons.

The following example further elucidates the relationship between self-efficacy and teacher motivation. In learning situations, motivation can be enhanced when individuals perceive they are making progress (Schunk, 1989). In turn, as they work on tasks and become more skillful, they maintain a sense of self-efficacy for performing well. This is particularly relevant to this study because the role of the teacher is important at all stages of the motivational process (Williams and Burden, 1997). Learners are influenced by their personal feelings towards their teachers, and by the way activities are presented in class. It makes sense that motivated teachers are more likely to behave in ways that contribute positively to learner motivation, such as developing learner autonomy, presenting activities in interesting ways, and giving effective feedback.

Teachers with positive feelings and beliefs about themselves, about their students, and about their teaching environments, are more likely to nurture positive beliefs and feelings in learners.

Next, Dörnyei and Ushioda (2013) listed competence (i.e., feeling efficacious and having a sense of accomplishment) as one of the basic conditions of intrinsic motivation. Furthermore, teacher education has traditionally taken a very one-sided approach by placing most emphasis on subject-matter training, accompanied by some (often rather limited) participatory experience in an instructional context that is supposed to provide the practical skills. Teacher-training programs do not include any awareness raising about how to manage groups (e.g., they do not cover the main principles of group dynamics and effective leadership strategies, and do not offer any training in interpersonal skills and conflict resolution). As a consequence, most newly qualified teachers are hit hard by the harsh reality of everyday classroom life that is often referred to as the “reality shock” (Veenman, 1984, p. 143). They are at a loss when something ‘goes wrong’ in the class, and because they lack any explicit skills in how to handle such inevitable crises, many of them change their original student-centered teaching behaviors and adopt a more authoritarian approach.

Finally, Dörnyei and Ushioda (2013) argue that newly qualified teachers are thrown into very deep water, and unless they have a natural ‘knack’ for dealing with people, they easily misunderstand and mishandle the inevitable fluctuation of both emotions and productivity that social groups regularly experience. Therefore, as a result of their lop-sided training, many teachers simply lack the skills necessary for doing well in the classroom. For them the task might be overly challenging and thus not intrinsically

motivating. This insufficient self-efficacy might very well contribute to the stress-generating nature of the teaching profession.

Self-efficacy and teacher anxiety.

Unfortunately, the relationship between self-efficacy and other motivational constructs is not always a beneficial relationship. In other words, high levels of certain affective phenomena like teacher anxiety, might lead to lower levels of self-efficacy. Bandura (1997) defined anxiety as “a state of anticipatory apprehension over possible deleterious happenings” (p. 137). Individuals experiencing anxiety embody uneasiness and avoidant behavior that often interferes with performance in everyday life. In social cognitive theory, one’s perceived sense of efficacy plays a key role in the arousal of anxiety. Those with a stronger sense of efficacy are more apt to take on the “deleterious happenings” that breed stress with positive expectations and are often more successful in transforming them into positive events. Individuals, therefore, only experience anxiety when they believe themselves to be incapable of managing potentially detrimental events (Bandura, 1997).

For example, Horwitz (1996), studied nonnative ESL teacher anxiety due to lack of self-confidence in the target language (English), and how this anxiety negatively influenced their class performance, thus contributing to anxiety and negative feelings in L2 learners. She found that motivation and ego-investment play substantial roles in foreign language anxiety. In other words, those people who sincerely want to learn or teach a language might be more likely to experience anxiety than those who have no personal stake in the effort. She also concluded that there must be a desire to

communicate well in order to worry about how your communicative efforts are being perceived in the first place. In her study, anxiety is related to self-efficacy and she ultimately shows the interaction and mutual influence between negative teacher feelings and negative feelings on the part of the learners.

Furthermore, in social cognitive theory (Bandura, 1997), self-efficacy plays a central role in the arousal of an individual's anxiety. Peoples' efficacy beliefs possess a positive relationship with their performance, whereas anxiety assumes a minimal or non-existent relationship to their achievement (Bandura, 1986, 1997). In other words, according to social cognitive theory, anxiety is typically a function of the lack of confidence with which individuals approach activities and tasks. As a result of a person's weakened sense of efficacy in a particular field, he or she becomes anxious about the corresponding demands (Mills, Pajares, & Herron, 2006). Consequently, anxiety serves as both a source and effect of self-efficacy beliefs.

Self-efficacy and teacher burnout.

Another negative yet common phenomenon in educational settings is teacher burnout.

Burnout is described as a psychological syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment that can occur among individuals who work with other people in some capacity. Emotional exhaustion refers to feelings of being emotionally overextended and depleted of one's emotional resources. Depersonalization refers to a negative, callous, or excessively detached response to other people, who are usually the recipients of one's services or care (Maslach, 1993, pp. 20-21).

Reduced personal accomplishment is described as a person's negative self-evaluation in relation to his or her job performance (Schaufeli, Maslach & Marek, 1993). The social

relationships with the people in professionals' care are the most obvious characteristic of occupations wherein burnout is an issue. Therefore, a social psychological perspective, such as self-efficacy, which incorporates social interactions, can be useful in acquiring a clearer understanding of burnout (Van Dierendonck, Schaufeli, & Sixma, 1994).

In several studies, teacher burnout has been shown to be moderately related to teacher self-efficacy (Chwalisz, Altmaier, & Russell, 1992; Evers, Brouwers, & Tomic, 2002; Friedman & Farber, 1992). More recently however, using structural equation modeling Skaalvik and Skaalvik (2010) found a strong relationship between teacher self-efficacy and teacher burnout. The connection was explored by means of a confirmatory factor analysis. They tested a model specifying three correlated latent variables; a second order teacher self-efficacy variable, and two primary burnout variables; emotional exhaustion and depersonalization. The model had acceptable fit to the data ($\chi^2(449, N = 2249) = 2345.76$, CFI = .949, and RMSEA = .043). Teacher self-efficacy correlated negatively with both emotional exhaustion (-.29) and depersonalization (-.41). The two dimensions of teacher burnout were positively, but weakly correlated (.23).

Finally, Walker and Symons (1997), in discussing social motivation theory and burnout, reported a phenomenon among teachers, where over time, they might develop negative feelings towards students. They explained that if teachers perceive negative feelings on the part of the students, and if they feel unappreciated in spite of their best efforts, then they respond with anger and anxiety, which can lead to low self-efficacy. They further elucidated that for these teachers, showing contempt for students is a way of preserving their self-esteem. Subsequently, self-efficacy theory has become a conceptual

framework that can be used to study the antecedents to teacher burnout (Brouwers & Tomic, 1998; Leiter, 1992).

Self-Efficacy and Gender

Bandura (1977) defined self-efficacy as the perceived ability of an individual to succeed at or accomplish certain tasks. Academic self-efficacy is essential to academic success (Lent & Hackett, 1987). The criterion-related validity of academic self-efficacy has been documented by several studies and one larger meta-analysis. Multon (1991) analyzed 36 studies that examined the relation of academic self-efficacy with performance and persistence and identified a mean correlation of $r = 0.38$ for performance and $r = 0.34$ for persistence. Given the importance of academic self-efficacy to academic achievement and persistence, determining whether academic self-efficacy is associated with other important variables, like gender, is worthwhile.

Gender differences in academic self-efficacy have been investigated extensively in recent decades. Although many researchers have examined gender differences in academic self-efficacy, findings have been inconsistent. As conventional narrative reviews (Pajares, 2002) can be influenced by subjectivity and bias, meta-analyses that quantitatively summarize studies are an alternative to a narrative review. In a meta-analysis of American and Canadian participants, Whitley (1997) examined gender differences in computer self-efficacy and found that the mean effect size was weak to moderate. The major limitation of this meta-analysis was its narrow scope, as it focused specifically on computer self-efficacy and ignored many relevant studies conducted outside North America. Generally, Eastern cultures emphasize collectivism, while

Western cultures prioritize individualism. Accordingly, Scholz (2002) explored cultural differences in self-efficacy by investigating whether general self-efficacy was a universal construct for 25 countries. He demonstrated that participants from collective cultures, such as those in Japan and Hong Kong, had low self-efficacy. As culture can play a role in the determination of academic self-efficacy, this meta-analysis is an extension of the meta-analysis by Whitley (1997) of gender differences in computer self-efficacy. Specifically, gender differences in all major components of academic self-efficacy were examined in individualistic and collective cultures.

As mentioned above, narrative reviews have been conducted regarding gender differences in academic settings. Pajares (2002), who summarized research on gender differences in math self-efficacy, reached four major conclusions. First, most studies indicated that male students had higher mathematics self-efficacy than females, while other studies did not. This inconsistency was related to variables used in regression equations. Second, gender differences in mathematics self-efficacy typically develop during middle school and increase as student age increases. Third, female students do not have higher mathematics self-efficacy than male students at any educational level. Finally, male students typically have higher mathematics self-efficacy than females, even when males and females have comparable achievement levels or when females outperform males. The pattern of gender differences in writing self-efficacy differs from that in mathematics self-efficacy. Pajares (2002) reviewed literature on gender differences in writing self-efficacy, concluding that females generally have higher writing self-efficacy than males during middle school; this gender gap disappears or reverses as the students become older. For gender differences in self-efficacy for self-regulated

learning, Pajares proposed that female students were generally more confident than male students.

In a meta-analysis of 82 studies with 104 effect sizes based on 40,491 American and Canadian participants, Whitley (1997) sought to elucidate gender differences in computer self-efficacy that yielded a mean effect size of $r = 0.41$, indicating that average computer self-efficacy for males was 0.41 standard deviations above the average computer self-efficacy for females. Significant heterogeneity among effect size estimates was a function of participant age. High school students had higher mean effect size than college and elementary school students. Furthermore, adult and college students had higher mean effect sizes than elementary school students. That is, high school students had higher computer self-efficacy than college students, who in turn had higher computer self-efficacy than elementary school students.

Finally, in a recent meta-analysis of 187 studies ($N = 68,429$) on gender differences in academic self-efficacy, Huang (2013) identified an overall effect size of 0.08, slightly favoring males. Moderator analysis demonstrated that content domain was a significant moderator in explaining effect size variation. Females displayed higher language arts self-efficacy than males. Meanwhile, males exhibited higher mathematics, computer, and social sciences self-efficacy than females. Gender differences in academic self-efficacy also varied with age. The largest effect size occurred for respondents aged over 23 years old. For mathematics self-efficacy, the significant gender differences emerged in late adolescence. Huang suggests that future research should longitudinally examine gender differences in academic self-efficacy to determine the prevalence of gender differences during different life stages.

Sources of Efficacy Beliefs

According to Bandura (1997), there are four main influences to a person's self-efficacy beliefs: mastery experiences, vicarious experiences, verbal persuasion, and physiological states. As Labone (2004) summarized:

Mastery experience refers to efficacy information gained from an individual's performance on a particular task. Vicarious experience refers to efficacy information gained from observing models perform a particular task including self-modeling. Verbal persuasion refers to efficacy information gained from positive talk about an individual's capability to perform a particular task. Physiological and affective states refer to efficacy information gained from physiological and emotive reactions to a particular task (p. 343).

The most influential source of information comes from mastery experiences because they provide the most realistic information to learners' concerning their ability to do whatever it takes to succeed. As learners' skills advance, their expectation that they will be able to master those skills further increases. Success tends to raise self-efficacy, whereas failure tends to lower it (Bandura, 1997; Driscoll, 2000). Furthermore, according to Tschannen-Moran et al. (1998), teachers' sense of efficacy is most directly influenced by mastery experiences and the emotional reactions associated with the experiences. This is because only in a situation of actual teaching can an individual assess the capabilities she or he brings to the task and experience the consequence of those capabilities.

The second source of efficacy beliefs is through vicarious experiences facilitated by modeling. Thus, modeling can serve as another tool for promoting self-efficacy. The more closely the observer identifies with the model, the stronger will be the impact on efficacy. Observing others perform tasks successfully raises expectations of personal success on the same task (Bandura, 1997). When it comes to teachers' sense of efficacy,

vicarious experiences and modeling are conceived as a powerful tool in pre-service teacher education (Labone, 2004; Tschannen-Moran et al., 1998).

The third source of self-efficacy is verbal persuasion. One's self-efficacy beliefs can be enhanced when others provide positive verbal judgments about their capabilities to perform certain tasks. This refers to others persuading a learner that he or she is capable of succeeding at a particular task (Driscoll, 2000). Bandura (1997) considers verbal persuasion to be a slightly weaker method of altering efficacy beliefs. While verbal persuasion is capable of influencing the learner to perform certain tasks, it tends to be disregarded by the learner if it is not verified to be true. In the context of teachers' sense of efficacy, Tschannen-Moran et al. (1998) stated, "verbal persuasion can be general or specific: it can provide information about the nature of teaching, give encouragement and strategies for overcoming situational obstacles, and provide specific feedback about a teacher's performance" (p. 219).

Finally, the fourth source of self-efficacy is physiological or affective states. This final source demonstrates how one's beliefs can be influenced by mood, stress level, anxiety, and subjective threats. For example, learners can stop performing a task because they tend to associate emotional arousal, such as anxiety or fear, as signs of personal incapability (Bandura, 1997). In the context of teaching, Tschannen-Moran et al. (1998) explained that high levels of arousal can impair functioning and interfere with making the best use of one's skills and capabilities, while moderate levels of arousal can improve performance by focusing attention and energy on the task.

Thus far, I have examined Bandura's social cognitive theory and self-efficacy beliefs. I first reviewed social cognitive theory and discussed its core features, including

human agency and the triadic reciprocal causation model. This model views human functioning as the product of a dynamic interplay between personal, behavioral, and social influences. Next, I examined the concept of self-efficacy and its place in social cognitive theory. Self-efficacy beliefs were shown to affect one's motivation level, effort/investment level, affective states, and accomplishments. Additionally, I reviewed not only how self-efficacy differs from other self-referent constructs, such as self-esteem and self-concept, but also how it is situated amongst various motivational variables, such as anxiety and burnout. The four sources of self-efficacy beliefs, where mastery experiences are the most powerful, were also examined. In the following section, I review literature on teachers' sense of efficacy.

Teachers' Self-Efficacy and its Measurement

In the following statement, Henson (2002) elucidated how the conception of teachers' sense of efficacy draws heavily on social cognitive theory:

Teacher efficacy as a construct has primarily stemmed from Bandura's (1997) social cognitive theory ... which suggests that one's efficacy beliefs are impacted by two important components: human agency and triadic reciprocal causation ... The interplay between these symbiotic influences results in actual behavior and thought in the individual. In this model, social context, perception, and behavioral action all impact a teacher's judgment about whether she or he will be able to execute the actions necessary to positively impact student learning. (p. 822)

The idea that teachers' beliefs about their own capabilities somehow matter (Tschannen-Moran, et al., 1998) has potentially powerful implications and has led to a variety of definitions. The definition of teacher self-efficacy adopted in this study is a "teacher's belief in his or her own capability to organize and execute courses of action required to

successfully accomplish a specific teaching task in a particular context” (Tschannen-Moran, et al., 1998, p. 233).

Teacher efficacy studies go back nearly four decades to the mid-1970s with RAND researchers’ examination of teacher characteristics and student learning (Armor, Conroy-Oseguera, Cox, King, McDonnell, Pascal, Pauly, & Zellman, 1976). Since then, numerous researchers have considered the meaning and attempted to measure this construct based on two theories: Rotter’s (1966) locus of control theory and Bandura’s (1977) social cognitive theory. In the following section, I examine various efforts to measure teacher efficacy based on both of the above theories. Finally, in response to conceptual confusion within teacher efficacy research, the most recent integrated model is introduced.

Measurement of Teacher Efficacy Based on Rotter’s Theory

Rotter’s theory of locus of control.

In social psychology, locus of control refers to the extent to which individuals believe that they can control events that affect them. This theory was developed by Julian Rotter (1954) and is central to his social learning theory of personality. Rotter (1966) later developed a 23-item forced choice scale, which has since become one of the most well-known questionnaires purporting to measure locus of control. Rotter explained that individuals with a high internal locus of control believe that events result primarily from their own behavior and actions. They are said to have better control of their behavior and are more likely to attempt to influence other people than those with a high external (or low internal) locus of control. On the other hand, individuals with a high external locus of

control believe that powerful others, fate, or chance primarily determine events. In contrast with those with a high internal locus of control, they are more likely to assume that their efforts will either be successful or not successful based on external factors.

Rotter's theory became quite well known by the 1970's when the concept of self-efficacy was first emerging in educational psychology. Incidentally, the first instruments purported to measure teacher self-efficacy were grounded in this theory. The following section highlights three studies that were developed based on Rotter's theory.

The Rand studies.

Studies by the RAND Corporation, funded by Title III of the Elementary and Secondary Education Act, are generally credited as being the earliest measures of teacher efficacy. Although current conceptions of teachers' sense of efficacy draw mainly on Bandura's social cognitive theory, as mentioned above, earlier studies in teacher efficacy were grounded in Rotter's (1966) theory of locus of control. The Rand Corporation conceptualized teacher efficacy as "the extent to which the teacher believes that he or she has the capacity to affect student performance" (McLaughlin & Marsh, 1978, p. 84).

In these studies, teachers were asked to respond to two 5-point Likert-type items ranging from strongly agree to strongly disagree. The two items used to measure teacher efficacy were: (a) "When it comes right down to it, a teacher really can't do much because most of a student's motivation and performance depends on his or her home environment," and (b) "If I really try hard, I can get through to even the most difficult or unmotivated students." These items were designed to measure the degree to which

teachers consider environmental factors as overwhelming any power they can exert in schools (external locus of control) or accept personal responsibility for what happens to them (internal locus of control), respectively (Guskey & Passaro, 1994). As a summed score of the two items, teacher efficacy was conceptualized as “teachers’ beliefs that factors under their control ultimately have greater impact on the results of teaching than factors in the environment or in the student-factors beyond the influence of teachers” (Tschannen-Moran et al., 1998, p. 205).

Although these measures provide important implications for teacher efficacy research, several researchers tried to expand the construct of teacher efficacy, and to develop longer and more comprehensive measures because of the reliability problems encountered with only two items.

Teacher locus of control.

Rose and Medway (1981) developed a 28-item forced-choice scale called the Teacher Locus of Control (TLC) that measures teachers’ tendencies to attribute student success and failure to an internal or external locus of control. An example from the scale is: “When the grades of your students improve, it is more likely (a) because you found ways to motivate the students, or (b) because the students were trying harder to do well.” The scale was found to be internally consistent and predictive of teacher behavior, such as willingness to implement new teaching strategies and the use of disciplinary techniques.

Webb efficacy scale.

Developed around the same time as the TLC, the Webb efficacy scale was an attempt to increase the reliability of the RAND measure while maintaining a narrow conceptualization of the construct (Ashton & Webb, 1986). The instrument consisted of seven items to which respondents were forced to choose whether they agreed more with either the first or second statement. For example, “(a) A teacher should not be expected to reach every child; some students are not going to make academic progress, or (b) Every child is reachable. It is a teacher’s obligation to see to it that every child makes academic progress.” Ashton and Webb (1986) reported that this measure had psychometric problems including inadequate reliability. Tschannen-Moran et al. (1998) reported that the Webb scale has not been widely used in the literature.

Measurement of Teacher Efficacy Based on Bandura’s Theory

Gibson and Dembo’s teacher efficacy scale (TES).

The Teacher Efficacy Scale developed by Gibson and Dembo (1984) is a 16-item questionnaire that uses a 6-point Likert scale ranging from Strongly Disagree to Strongly Agree. It has been one of the most widely used scales purporting to measure self-efficacy. Through a factor analysis of 208 elementary teachers’ responses, they reported a two-factor model that accounted for 28.8% of the total variance. Gibson and Dembo noted that Factor 1 represents a teacher’s sense of personal teaching efficacy, and corresponds to Bandura’s self-efficacy dimension. On the other hand, the second dimension stands for the teacher’s sense of teaching efficacy, and corresponds to Bandura’s outcome expectancy dimension. They called these dimensions Personal Teaching Efficacy and

General Teaching Efficacy, respectively. In other words, general teaching efficacy referred to the teachers' general beliefs about the possibility of producing student learning in the face of multiple obstacles (e.g. unsupportive home environment) and personal teaching efficacy referred to the teacher's personal appraisal of his or her own effectiveness as a pedagogue. They presented internal consistency reliability alpha coefficients of .78 for Personal Teaching Efficacy, .75 for General Teaching Efficacy, and .79 for all 16 items on a short form of the scale. (Note: The long form of the TES contains 30 items.) They recommended the use of the revised 16-item scale for further research.

Guskey and Passaro (1994) noticed that there were biases in the wording of the items on the TES. Items measuring personal efficacy used the referent "I" and were positive; while items measuring teaching efficacy used "teachers" and were negative. For that reason, they changed the wording of the items in order to have balanced characteristics throughout the instrument. They administered this scale to a sample of 342 teachers, and reported that the results confirmed internal and external dimensions instead of personal and teaching efficacy dimensions. They reported:

The internal factor appears to represent perceptions of *personal* influence, power, and impact in teaching and learning situations. Because of the nature of the items in the current scale, these perceptions reflect a perspective that is positive and optimistic. The external factor, on the other hand, relates to perceptions of the influence, power, and impact of elements that lie *outside the classroom* and, hence, may be beyond the direct control of individual teachers. (p. 639)

Conceptual confusions.

Although the Gibson and Dembo instrument has been widely used and adapted, conceptual and statistical problems remain unresolved (Tschannen-Moran et al., 1998).

As explained above, due to concerns about construct clarity and lack of measurement reliability with the two items developed by the RAND researchers (1976), Gibson and Dembo (1984) developed a new instrument to measure teacher efficacy. They argued that the two items used by the RAND researchers based on Rotter's (1966) theory actually corresponded to Bandura's (1977) outcome expectancy and self-efficacy dimensions of social cognitive theory, respectively.

Bandura's idea about outcome expectancy is theoretically independent from self-efficacy. As explained earlier, an efficacy expectation is one's belief that he or she can perform the necessary actions to complete a given task, while outcome expectancy is one's estimate of the likely consequences of performing that task at the expected level of competence (Bandura, 1986). In contrast to an efficacy question, 'Do I have the ability to organize and execute the actions necessary to accomplish a specific task at a desired level?' an outcome question asks, 'If I accomplish a task at this level, what are the likely consequences?'

RAND Item 1 was thought to assess an outcome expectancy regarding a teacher's belief about whether teaching can impact student learning despite external constraints. This construct was labeled Teaching Efficacy (later to be called General Teaching Efficacy, or [GTE]). RAND Item 2 was thought to assess self-efficacy, or a teacher's perceived ability to positively impact student learning. This construct was named Personal Teaching Efficacy (PTE). As mentioned above, Gibson and Dembo (1984) developed their items based on the RAND items and after a factor analysis, retained 16 items on the short form of the scale (PTE, 9 items; GTE, 7 items). The PTE and GTE factors were negatively correlated ($r = -.19$), a result consistent with Bandura's

conceptualization of the independence of the outcome expectancy and self-efficacy dimensions. To demonstrate construct validity, Gibson and Dembo (1984) then performed a multitrait–multimethod study that resulted in the Teacher Efficacy Scale (TES). The TES quickly became the leading instrument in the study of teacher efficacy (Ross, 1994).

As teacher efficacy research grew, serious questions about the TES arose. Specifically, in a study of correlations among scores from the major instruments measuring teacher efficacy and related constructs, Coladarci and Fink (1995) found weak evidence for the discriminant validity of PTE and GTE scores. Furthermore, Guskey and Passaro (1994) reported that the PTE and GTE factors corresponded not to self-efficacy and outcome expectancy dimensions, but to an internal versus external orientation, respectively. This split resembled locus of control and attribution theory orientations more than self-efficacy theory (Henson, 2002). Especially noteworthy is that the studies by Coladarci and Fink (1995) and Guskey and Passaro (1994) displayed potential theoretical confounds in the TES.

The single most critical matter in the conceptual confusion surrounding self-efficacy instruments is that the TES was originally developed from the two RAND items, which were based on locus of control theory (Rotter, 1966). Gibson and Dembo (1984) then interpreted the items as reflecting self-efficacy theory (Bandura, 1977). Consequently, the TES appears to have elements of both theoretical orientations captured in its items. Henson (2002) explained that as might be expected from an instrument that serves two theoretical masters, the study of teacher efficacy has suffered an adolescent identity crisis as researchers have struggled to clarify the self-efficacy construct.

Bandura's teacher self-efficacy scale.

In an attempt to clarify some of the confusion, Bandura developed his own teacher efficacy scale, which is a 30-item instrument with seven subscales: efficacy to influence decision making, efficacy to influence school resources, instructional efficacy, disciplinary efficacy, efficacy to enlist parental involvement, efficacy to enlist community involvement, and efficacy to create a positive school climate. Each item is measured on a 9-point scale anchored by the following descriptors: *Nothing* = 1, *Very little* = 3, *Some influence* = 5, *Quite a bit* = 7, *A great deal* = 9 (Bandura, 2001). By presenting a broader framework of teaching, this scale represents an improvement over previous measures of teacher efficacy. The consideration of factors outside the classroom, such as a teacher's ability to influence policy and social interaction, provide greater insight into the possible dimensions of teacher self-efficacy. While Bandura's scale offered considerable augmentation of the evaluative framework stemming from the field of teaching, greatly influencing later scales such as the TSES, the need still remained to include domain-specific items across multiple contexts of instruction, such as TESOL education. Additionally, reliability and validity information is not available for this scale, and it should be used cautiously without this information. Moreover, because Bandura's scale was not ultimately published until 2007 after spending many years as an 'underground' scale only familiar to those self-efficacy researchers connected to him, as well as being published among many other sample instruments in a guide to creating self-efficacy scales, this instrument is not widely known and has hardly been used in published studies. Finally, the generality of the scale is not specific enough and likely

either over-represents or under-represents the individual teaching tasks and environments for many teachers.

Table 1 provides an outline of six instruments that have been used to measure teacher self-efficacy. The first three instruments are based on Rotter's (1966) locus of control theory and the last three instruments are based on Bandura's (1977) self-efficacy theory. The table provides information about the types of scales used and shows example items.

Table 1. Teacher Self-Efficacy Measures, Scale Types, and Example Items

Measure	Scale type	Example item
RAND measure (Armor et al., 1976)	Two items on a 5-point Likert scale ranging from <i>Strongly agree</i> to <i>Strongly disagree</i>	When it comes right down to it, a teacher really can't do much because most of a student's motivation and performance depends on his or her home environment.
Teacher Locus of Control (Rose & Medway, 1981)	28 items with an A/B forced-choice format	If you could not keep your class quiet, it would probably be (a) because the students came to school more rowdy than usual, or (b) because you were so frustrated that you were not able to settle them down?
Webb Efficacy Scale (Ashton & Webb, 1986)	Seven items with an A/B forced choice format	(a) My skills are best suited for dealing with students who have low motivation and who have a history of misbehavior in school. (b) My skills are best suited for dealing with students who are academically motivated and generally well behaved.
Teacher Efficacy Scale (Gibson & Dembo, 1984)	30 items on a 5-point Likert scale ranging from <i>Strongly agree</i> to <i>Strongly disagree</i>	If a student masters a new concept quickly it is usually because I knew the necessary steps in teaching that concept.
Bandura's Teacher Efficacy Scale (Bandura, 2001)	30 items on a 9-point Likert scale ranging from <i>Nothing</i> to <i>A great deal</i>	How much can you do to make students enjoy coming to school?
Teacher Sense of Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001)	24 items on a 9-point Likert scale ranging from <i>Nothing</i> to <i>A great deal</i>	How much can you do to calm a student who is disruptive or noisy?

An Integrated Model

The construct of teacher efficacy suffered from the conceptual confusions of Rotter's locus of control theory and Bandura's social cognitive theory, as indicated by the confounds discovered repeatedly through testing the Gibson and Dembo (1984) instrument. For the purpose of addressing the confusions and bringing coherence to the meaning and measure of teacher efficacy, Tschannen-Moran et al. (1998) proposed an integrated model (see Figure 2).

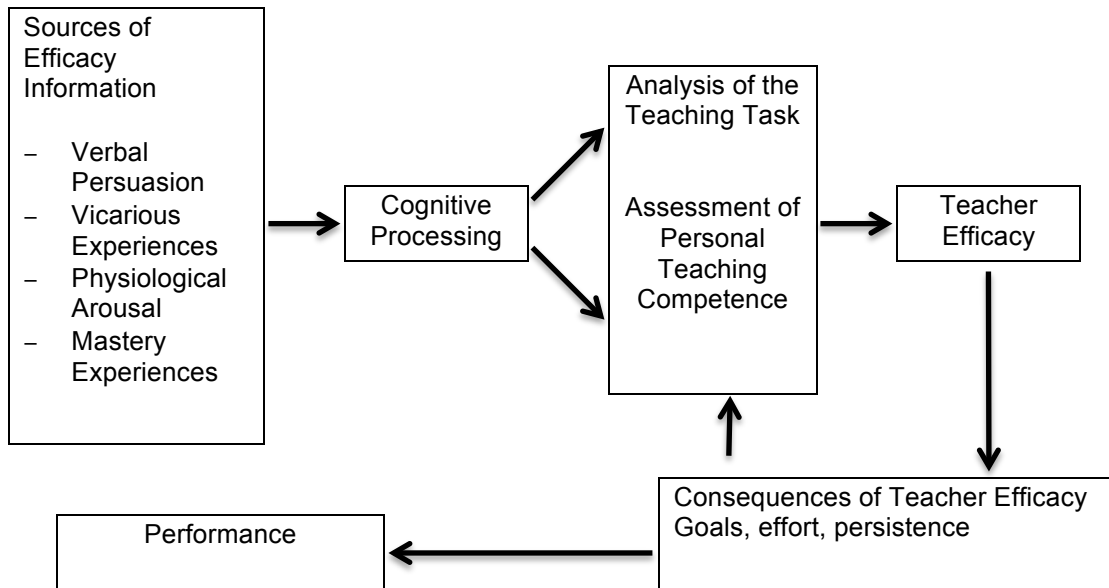


Figure 2. Teachers' Sense of Efficacy Integrated Model (Tschannen-Moran et al., 1998).

As shown in Figure 2, the model is based largely on Bandura's self-efficacy theory in terms of sources of efficacy information, cognitive processing, the domain-specific nature of self-efficacy beliefs (analysis of the teaching task), consequences of

teacher efficacy, and most notably, the cyclical nature of self-efficacy beliefs. In addition, this model points to an analysis of teaching tasks and does not just focus on the constraints facing teachers in general (as conceived in General Teaching Efficacy) in conceptualizing teacher efficacy. Thus, in this model, teachers weigh their self-perceptions of personal teaching competence in light of the assumed requirements of the anticipated teaching task when making judgments of self-efficacy.

Teacher efficacy, in this model, is defined as “the teachers’ belief in his or her capability to organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context” (Tschannen-Moran et al., 1998, p. 222). In this conception, what is particularly relevant is the nature of the domain and context specificity of teacher efficacy. In relation to this, Bandura (1997) rejected most of the existing teacher efficacy scales because they were too general instead of being tailored to domains of instructional functioning. Pajares (1996) also noted:

Omnibus tests that aim to assess general self-efficacy provide global scores that decontextualize the self-efficacy—behavior correspondence and transform self-efficacy beliefs into a generalized personality trait rather than the context-specific judgment Bandura suggests they are... The problem with such assessments is that students must generate judgments about their academic capabilities without a clear activity or task in mind. As a result, they generate the judgments by in some fashion mentally aggregating to related perceptions that they hope will be related to imagined tasks. (p. 547)

Clearly, one of the most challenging tasks to date has been for self-efficacy researchers to locate the ideal balance between generality and specificity within their instruments. On the one hand, it is necessary to make the items context-specific enough to avoid broad generalizations. However, on the other hand, an item that is too specific obstructs any generalization at all.

Teachers' Sense of Efficacy Scale (TSES)

Taking the Bandura teacher efficacy scale as a base and using their new integrated teachers' sense of efficacy model, Tschannen-Moran and Woolfolk Hoy developed the TSES, previously called the Ohio State Teacher Efficacy Scale, in a seminar on self-efficacy in teaching and learning at the Ohio State University. The seminar participants searched for an instrument, which included the types of tasks representative of frequent teaching activities. They worked to develop and add new items and they decided to use a 9-point scale as Bandura did. Tschannen-Moran and her colleagues then investigated the resulting instrument in different studies.

In the initial study, the 52-item instrument was administered to a sample of 224 pre-service and in-service teachers. Thirty-two of the items believed to measure three constructs, Student Engagement, Instructional Strategies, and Classroom Management, were selected as a result of principal-axis factoring with varimax rotation (Tschannen-Moran & Woolfolk Hoy, 2001).

In the second study, the 32-item version of the TSES was investigated with a sample of 70 pre-service and 147 in-service teachers. Tschannen-Moran and Woolfolk Hoy (2001) used principal axis factor extraction again. The eigenvalue greater than one criterion yielded an eight-factor solution, while the scree test suggested a possible two- or three-factor solution. After examining both two- and three-factor solutions, the authors chose the three-factor solution, of Student Engagement, Instructional Strategies, and Classroom Management, which seemed to represent typical teaching tasks.

Later, the instrument was reduced to 18 items by removing redundant items and items with low factor loadings. The factor analysis with varimax rotation produced three

factors accounting for 51% of the variance. These factors were called Efficacy for Student Engagement (eight items with an alpha reliability of .82), Efficacy for Instructional Strategies (seven items with an alpha reliability of .81), and Efficacy for Classroom Management (three items with an alpha reliability of .72). A further analysis, using samples from study 1 and study 2, generated one strong factor with factor loadings ranging from .74 to .84. Tschannen-Moran and Woolfolk Hoy (2001) argued that the TSES could be used for the assessment of either three domains of efficacy or one generalized efficacy factor.

Roberts and Henson (2001) evaluated the 18-item instrument while expressing their concerns about a third factor that was measured with only three items. They collected data from 183 in-service teachers, and subjected the data to a confirmatory factor analysis (CFA). The CFA approach differed from Tschannen-Moran and Woolfolk Hoy (2001), who had used an exploratory factor analysis. Their findings supported the factorial validity of the TSES but for only the Efficacy for Student Engagement and Efficacy for Instructional Strategies factors. Roberts and Henson (2001) argued that the items measuring the third factor, Efficacy for Classroom Management, should be removed from the instrument because the items had low loadings. In addition, they rejected the unidimensional model suggested by Tschannen-Moran and Woolfolk Hoy (2001) because it contradicted the CFA results.

On the other hand, Tschannen-Moran and Woolfolk Hoy (2001) argued that classroom management is a crucial factor in teaching and disagreed with the elimination of this factor. Instead, they developed new items concerning classroom management by taking Emmer's (1990) teacher efficacy for classroom management scale into

consideration. The resulting instrument included 36 items. After administering the instrument to a sample of 410 pre-service and in-service teachers, Tschannen-Moran and Woolfolk Hoy then employed principal-axis factoring with varimax rotation, resulting in a four-factor solution suggested by the eigenvalues greater than 1.0 criterion, and a three-factor solution suggested by the scree test. The three-factor solution was consistent with the findings of study 2.

Lastly, Tschannen-Moran and Woolfolk Hoy (2001) selected items with the highest loadings and developed a short form (12 items) and long form (24 items) of the instrument. Analyses of both forms indicated that both the short and long versions of the TSES were reliable and valid instruments for assessing the teacher efficacy construct. Both versions supported the three-factor model with high subscale reliabilities ranging from .87 to .91 for the longer version and .81 to .86 for the shorter version.

Furthermore, Tschannen-Moran and Woolfolk Hoy (2001) examined the validity of the long and short versions of the TSES by assessing the relationship of the instruments with previously used instruments, including the Rand items and the 10-item adaptation of Gibson and Dembo's Teaching Efficacy Scale. Consistent with the previous studies, the most significant parallels among measures were found to be with the scales assessing personal teaching efficacy rather than general teaching efficacy.

To summarize the development of the TSES, Tschannen-Moran, Woolfolk Hoy, and eight graduate students started with Bandura's (1997) teacher efficacy questionnaire and developed a 52-item instrument that used a 9-point scale. This questionnaire was examined in three studies. In the first two studies, the original 52 items were reduced to 32 and then to 18 items. In the third study, 18 additional items were developed and tested,

resulting in two forms of scale: 24-item and 12-item scales. The three studies consistently produced three factors: Classroom Management, Instructional Strategies, and Student Engagement (Tschannen-Moran & Woolfolk Hoy, 2001).

I have so far reviewed literature on teachers' sense of efficacy, focusing on its conception and measurement and the factors associated with teachers' sense of efficacy. The history of studies on teachers' sense of efficacy shows that the construct of teacher efficacy has suffered from confusion of the two conceptual strands, Rotter's locus of control theory and Bandura's social cognitive theory. It is clear there have been an abundance of problems with the measurement of teachers' sense of efficacy. In a comprehensive review of literature on teachers' sense of efficacy, Tschannen-Moran, et al. (1998) proposed a new integrated model of teacher efficacy (see Figure 3) in which one's judgment of personal teaching competence is identified through specific teaching tasks. In this conception of teacher efficacy, the domain and context specificity and the cyclical nature of teacher efficacy are highlighted. This model has been acknowledged as a significant advancement in teacher efficacy research (Henson, 2002; Labone, 2004). Based on this integrated model, the TSES was developed in 2001 and has been used in recent research (Klassen, 2009; Skaalvik, 2010). In accordance with this, in this study I modify the TSES to fit the specific teaching context in question, the university English class setting in Japan. In the next section, I first review studies attempting to create context- and subject- matter specific scales. I then review the few studies on teachers' sense of efficacy in the TESOL field.

Subject-Matter and Context-Specific Scales

Teacher efficacy has been defined as both context- and subject matter-specific. A teacher might feel competent in one area of study or when working with one kind of student and feel less able in other subjects or with different students. While researchers and theorists agree that teacher efficacy is situation specific, it is less clear what the appropriate level of specificity is (Tschannen-Moran, et al., 1998). For example, is efficacy specific to teaching English, or more specific to British literature, or even more specific to teaching the intricacies of William Shakespeare's writing? Recognizing that many standard efficacy instruments overlook the specific teaching context, some researchers have modified the Gibson and Dembo instrument to explore teachers' sense of efficacy within particular domain-specific areas.

Science teaching.

Science educators have conducted extensive research on the effects of efficacy on science teaching and learning. Riggs and Enochs (1990) developed an instrument, based on the Gibson and Dembo approach, to measure efficacy of teaching science, the Science Teaching Efficacy Belief Instrument (STEBI). Consistent with Gibson and Dembo, they identified two uncorrelated factors, one they called Personal Science Teaching Efficacy (PSTE) and a second they labeled Science Teaching Outcome Expectancy (STOE).

Teachers with a higher sense of PSTE, as measured using the STEBI, reported spending more time teaching science and were more likely to spend an ample amount of time to develop the science concept being considered (Riggs & Jesunathadas, 1993). Teachers with low PSTE spent less time teaching science, used a text-based approach,

were rated as weak by site observers, made fewer positive changes in their beliefs about how children learn science, and were less likely to choose to teach science (Riggs, 1995).

Teachers with low scores on the STOE were rated as less effective in science teaching, rated themselves as average, and were rated as poor in attitude by site observers (Enochs et al, 1995). Low-scoring teachers used text-based approaches over hands-on, activity-based approaches and used cooperative learning less (Riggs, 1995).

Teachers who were engaged in a year-long training program in science education (the Science Education and Equity Project) perceived differential changes in their efficacy beliefs, depending on their initial efficacy scores. Teachers who began the training with low scores on both scales of the STEBI made great gains in personal efficacy (PSTE) during the training, but outcome expectancy (STOE) remained the same. High efficacy-low outcome expectancy teachers had increases in both scales. Low efficacy-high outcome expectancy teachers increased in self-efficacy, but remained stable on outcome expectancy (Riggs, 1995). As the teachers implemented methods they learned in the training, they saw improved student achievement, and their personal teaching efficacy improved. The training improved the teacher's beliefs about what science teachers in general could achieve.

Special education.

To explore teacher efficacy in the context of special education, Coladarci and Breton (1997) used a 30-item instrument, modified from Gibson and Dembo (1984) and reworded to apply specifically to special education. Higher efficacy was found among teachers with high satisfaction, among women, and among teachers who were older;

however, the length of time a teacher had spent working in a resource room was not related to efficacy beliefs. In order to study the likelihood of referral to special education in the Netherlands, Meijer and Foster (1988) developed the Dutch Teacher Self-Efficacy Scales, an 11-item instrument examining personal teaching efficacy beliefs. Teachers were asked to respond along a 4-point Likert scale to statements such as “I become truly discouraged when I see a pupil returning to problem behavior” or “I can handle virtually any learning problem well.” The researchers found that high efficacy teachers were more likely to feel that a problem student was appropriately placed in the regular classroom.

Classroom management.

In 1990, Emmer and Hickman adapted the Gibson and Dembo instrument in an attempt to better reflect the domain of classroom management. This effort yielded a 36-item measure with three efficacy subscales: Efficacy for Classroom Management and Discipline, External Influences, and Personal Teaching Efficacy. The efficacy subscales were correlated with strategies aimed at increasing desirable student responses through praise, encouragement, attention, and rewards among a sample of pre-service teachers. The subscales were not related to attempts to limit or eliminate behaviors by using timeouts, punishment, or reprimands. Pre-service teachers with a higher sense of Personal Teaching Efficacy were more likely to seek outside help in dealing with student discipline problems (Emmer, 1990; Emmer & Hickman, 1990).

English as a Foreign Language Teachers' Self-Efficacy

In the TESOL field, research on teachers' sense of efficacy is scarce (Chacón, 2002; Shim, 2001). An Internet database search for this subject (EFL and/or TESOL and/or university teacher self-efficacy, sense of efficacy, or self-efficacy beliefs) yielded very few results. A few of the more substantial studies included two unpublished doctoral dissertations (Chacón, 2002; Shim, 2001), one international journal article (Chacón, 2005), and two Korean domestic journal articles (Shim, 2003, 2006). All of the studies were conducted in an English as a Foreign Language (EFL) setting, where English is not used for daily communication purposes.

TESOL and Self-Efficacy

Shim (2001) examined Korean public middle and high school English teachers' sense of efficacy with 108 teachers who responded to a 10-item questionnaire. Shim used Hoy and Woolfolk's (1993) version of Gibson and Dembo's (1984) Teacher Efficacy Scale (TES) to measure the teachers' sense of efficacy in relation to intrapersonal variables (teaching satisfaction, role preparedness, classroom management, and language skills), interpersonal variables (working relationship with peer teachers, perceptions of school stress, and level of academic emphasis), and demographic variables (gender, educational level, student grade level, and the experience of visiting English speaking countries).

Shim (2001) found that his participants had the same dimensions of efficacy as the American teachers identified in Gibson and Dembo (1984). The dimensions were Personal Teaching Efficacy (PTE), referring to the "belief that one has the skills and

abilities to bring about student learning” (p. 98), and General Teaching Efficacy (GTE), referring to the “belief that any teacher’s ability to bring out change is significantly limited by factors external to the teacher” (p. 98). Also, Shim reported that PTE had statistically significant correlations with such variables as teaching satisfaction, job preparedness, classroom management, academic emphasis, and listening skills, but GTE had no statistically significant correlation with any of these variables. He also found that the participants’ GTE was low ($M = 2.92$) indicating, “they did not expect their teaching to have an effect on their students’ learning” (p. 103). Shim suggested that future researchers identify the reasons behind the low score and help teachers enhance their GTE. Furthermore, Shim did not ask about teacher efficacy in relation to a specific subject matter, although his participants were public secondary English teachers. The global nature of the 10-item TES scale (Hoy & Woolfolk Hoy, 1993) might not have been specific enough for the English teachers to recall the typical tasks they were supposed to perform (cf., Bandura, 1997; Tschannen-Moran, et al., 1998; Tschannen-Moran & Woolfolk Hoy, 2001).

Unlike Shim’s study, Chacón (2002) took a subject matter specific approach. She examined 104 Venezuelan teachers’ sense of efficacy in teaching English by modifying different teacher efficacy scales (i.e., Teachers’ Sense of Efficacy Scale [Tschannen-Moran and Woolfolk Hoy, 2001]; Science Teacher Efficacy Scale [Riggs and Enochs, 1990]; and the two Rand items). She also added scales for job preparedness, proficiency in the four language skills (i.e., reading, writing, listening, reading), and instructional and management strategies. All of the private and public middle school English teachers were working in the same town in Venezuela at the time of the study, an aspect of the study

that offered little geographic diversity. She examined the teachers' sense of efficacy in relation to their self-reported language proficiency, self-reported teaching strategies, student management strategies, job preparedness, and professional backgrounds. She also conducted interviews with 20 participants.

Chacón (2002) found that the Venezuelan English teachers' sense of efficacy was correlated with their job preparedness, which supported Shim's (2001) finding. However, unlike Shim, Chacón (2002, 2005) found that the teachers' sense of efficacy was positively correlated with their five language skills (i.e., reading, writing, listening, reading, and culture). "Substantial positive correlations" (p. 114) were found between PTE and perceived English language proficiency with reading ($r = .51$), writing ($r = .55$), listening ($r = .56$), speaking ($r = .60$), and culture ($r = .56$) (p. 114). On the other hand, the correlations between GTE and language proficiency were positive but lower $r = .25$ to $r = .39$. Also, English language proficiency was found to have positive correlations with two of the TSES subscales (student engagement: $r = .26$ to $r = .34$, and instructional strategies: $r = .33$ to $r = .41$) but to have low correlation with classroom management, with the exception of writing ($r = .23$) (Chacón, 2005). Based on these findings, Chacón argued that for non-native speakers, English language proficiency is a strong predictor of teachers' sense of efficacy in teaching English. This finding was not consistent with Shim's 2001 results.

Shim (2006) tried to identify the dimensions of teacher efficacy beliefs specific to Korean secondary school teachers of English by surveying 118 teachers. A principal component analysis and confirmatory factor analysis identified four dimensions, Teaching Improvement, Classroom Management, Confidence in English, and Students'

Affect. Although Shim's dimensions provide a general picture about teachers' beliefs about themselves as English teachers, some of the items do not seem to be in accordance with the current theorization of teacher efficacy (e.g., context specificity) (Bandura, 1997; Tschannen-Moran, et al., 1998). For example, one of the items was "*I have low confidence in English,*" which does not ask about teacher's beliefs about their capability to teach English.

In this chapter I have reviewed studies on teacher efficacy in the TESOL field. The teacher-efficacy research is scarce, and only a few studies were found relevant to this area of inquiry. Given the powerful impact of teachers' sense of efficacy on various aspects of teaching and learning, this scarcity showed that there is a need to inquire into teacher's efficacy in the TESOL field, particularly in the university setting in Japan.

In the next section, I summarize the gaps identified in the literature review and present the research questions that guide this study.

Gaps in the Literature

Self-efficacy is defined and measured as specific to behaviors in specific contexts or situations (Bandura, 1997). Consequently, Bandura cautioned researchers examining self-efficacy beliefs that they should use assessments that correspond to the specific domain of functioning being analyzed. Otherwise, the resulting omnibus-type instrument will be unclear about what is being assessed. However, as can be seen in the review of the literature, in educational research it is not uncommon to find global measures of efficacy that decontextualize self-efficacy behavior correspondence and transform self-efficacy beliefs into a generalized personality trait (Pajares, 1996). Many researchers

have simply measured general self-perceptions of competence. Pajares argued that in these studies, respondents are made to assess their capabilities without a clear domain or task in mind. As a result, these assessments are not domain- and task-specific assessments, as efficacy judgments should be. Pajares also cautioned researchers to carefully consider the level of specificity so that they can generalize the findings and provide results that have practical utility. Pajares stated,

Judgments of competence need not be so microscopically operationalized that their assessment loses all sense of practical utility. Domain specificity should not be misconstrued as extreme situational specificity, and there is no need to reduce efficacy assessments to atomistic proportions. (p. 13)

It is clear that of the unresolved issues in the measurement of teacher efficacy, making the instrument domain-specific and determining the optimal level of specificity, is essential.

Furthermore, the literature reviewed in this chapter on teachers' self-efficacy beliefs in the TESOL field, while helpful, demonstrates that this line of research is scarce. Given its strong relationships with various aspects of teaching and learning (Henson, 2002; Labone, 2004; Tschannen-Moran & Woolfolk Hoy, 2001; Tschannen-Moran, et al., 1998; Woolfolk Hoy, et al., 2006), teacher efficacy deserves more attention in the EFL domain, specifically at the university level and in the Japanese context, where no literature currently exists.

Inconsistent relationships found in the studies of teachers' sense of efficacy highlight a need to further examine this relationship. This need seems to be especially pertinent to countries such as Japan, where English is widely taught in universities by thousands of native and non-native speakers, and assumptions concerning obstacles for

successful university teaching of English remain unidentified because of a lack of empirical data.

The lack of teacher efficacy research in the TESOL field also highlights a need to explore other factors that might be related to teachers' sense of efficacy in teaching English in different contexts and levels. As Shim (2003) pointed out, while language proficiency of non-native speakers explains some degree of variance (20% in his study) in teachers' sense of efficacy, the low percentage indicates that there are other factors in addition to English language proficiency that should be explored. Studies examining additional influential factors will inform English teacher educators and policy makers of the areas that need to be addressed in teacher education programs for English teaching. Finally, none of the studies reviewed were carried out in a university setting, none were conducted in Japan, and none included native English speakers.

The review of previous research has shown the following gaps:

1. Teacher self-efficacy researchers have utilized various scales and inventories to produce equivocal and widely diverse results, due mainly to a lack of measurement integrity and unclear theoretical perspectives.
2. Few studies have been conducted on the self-efficacy beliefs of EFL teachers.
3. To the best of my knowledge, no researchers have investigated the self-efficacy beliefs of Japanese university English language teachers while taking in to consideration the context-specific nature of an EFL teacher's work environment.

Purposes of the Study

Based on the gaps found in the literature, there are two purposes to this study. The first purpose, which addresses the first gap, is to introduce and validate an instrument appropriate for investigating Japanese university English teachers' self-efficacy beliefs. This is significant because no instrument currently exists that fully captures the intricacies of the jobs of university English teachers. The development and validation of such an instrument fills a void that currently exists in the field of TESOL and adds another valuable tool to the self-efficacy research repertoire.

The second purpose, which addresses the second and third gaps simultaneously, is to explore the self-efficacy beliefs of university English teachers in Japan using a mixed-methods approach. This is done in two ways. First, because this study was a baseline study, only major independent variables were considered for measurement. In other words, smaller variables such as classroom size and university major were not included. The final four variables chosen for measurement in this study came to be after several brainstorming sessions, recent literature reviews, and the input and critique of four professional university English teachers, currently working in Japan. I explore self-efficacy using a new robust instrument to compare and contrast various groups of university English teachers' self-efficacy levels. Specifically, I investigate the (a) teachers' native language (b) years of teaching experience, (c) tenured vs. contract status, and, (d) gender, to determine if the groups have parallel scores on the self-efficacy measure. Second, I conduct qualitative interviews to explore teacher supports (or lack of) that influence and challenge the development of teachers' self-efficacy beliefs.

Given the distinctiveness of the academic and administrative environment at

Japanese universities and the important roles the English teachers play in tertiary education, it is important to investigate the levels of teachers' self-efficacy to provide at the very least, a baseline picture of the current status of university English education in Japan from the teachers' perspective.

Research Questions

The main purpose of this study is to investigate Japanese university English language teachers' self-efficacy beliefs. The following research questions have been posited:

1. To what extent do Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' native language?
2. To what extent do Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' length of teaching experience?
3. To what extent do Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' tenure track vs. term-limited contract status?
4. To what extent do Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' gender?
5. What potentially supports strong language teacher self-efficacy beliefs?
6. What potentially challenges language teacher self-efficacy beliefs?

CHAPTER 3

METHOD

The purpose of this chapter is to explain the use of a mixed-method design, describe the participants, instrumentation, procedures for data collection, and the analyses used in this study.

Mixed-Method Research Design

In this study, I use a triangulation strategy (Creswell, 2009, p. 213), a mixed-method design in which the researcher collects both quantitative and qualitative data. This type of design enables the researcher to use results from one method to support and inform the other method. Furthermore, it allows the researcher to compare the two results to determine if there is convergence, differences, or some combination. Creswell explained that this type of design generally uses separate quantitative and qualitative methods as a means to offset the weaknesses inherent within one method with the strengths of the other. The data collection for both the quantitative and the qualitative can happen concurrently, sequentially, or a bit of both. Although the weight is ideally equal between the two methods, often in practice, priority is given to one or the other. While quantitative data which is characteristically represented by numbers or score totals on scales, is valuable and can produce clear overarching pictures about the phenomena being investigated, qualitative data obtained from open-ended interviews, for example, can “offer many different perspectives on the study topic and provide a more complex picture of the situation” (Creswell, 2005, p. 510). This point is particularly applicable to the

present study, as Pajares (1992) stated that when considering teachers' beliefs "additional measures such as open-ended interviews...must be included if richer and more accurate interpretations are to be made" (p. 327). Therefore, in order to minimize the limitations of the questionnaire used in this study, qualitative data consisting of open-ended interviews are used to help triangulate and better contextualize the quantitative data. The following Figure is a visual representation of the triangulation strategy employed in this study (Creswell, 2009, p. 210).

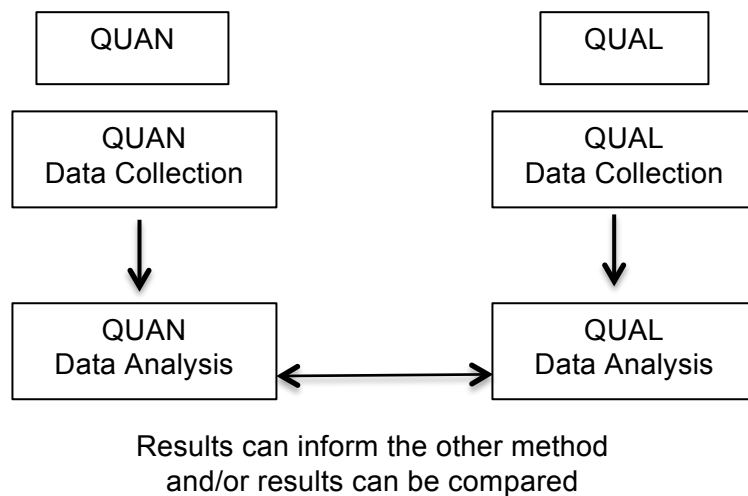


Figure 3. Triangulation strategy design (Creswell, 2009).

To summarize, this study uses a questionnaire to measure the participants' English teaching self-efficacy beliefs and the relationship these beliefs have with the participants' native language, teaching experience, tenure vs. contract status, and gender. The complementary qualitative part of the study consists of interviews exploring how teachers think their own teaching self-efficacy beliefs are created, shaped, and challenged.

Participants

Questionnaire

The respondents to the questionnaire were 440 English teachers currently working at Japanese universities. A combination of an exponential non-discriminative snowball sampling method (see Figure 5) as well as a stratified purposeful sampling method was adopted for this study. Snowball sampling is a non-probability sampling technique that is used by researchers to identify potential participants in studies where they are difficult to locate (Castillo, 2009). Researchers use this sampling method if the sample for the study is very rare or is limited to a small subgroup of the population. The researcher asks for assistance from the participants to help identify people with a similar trait of interest, in this case Japanese university English language teachers. The researcher then observes the nominated subjects and continues in the same way until a sufficient number of participants have been obtained.

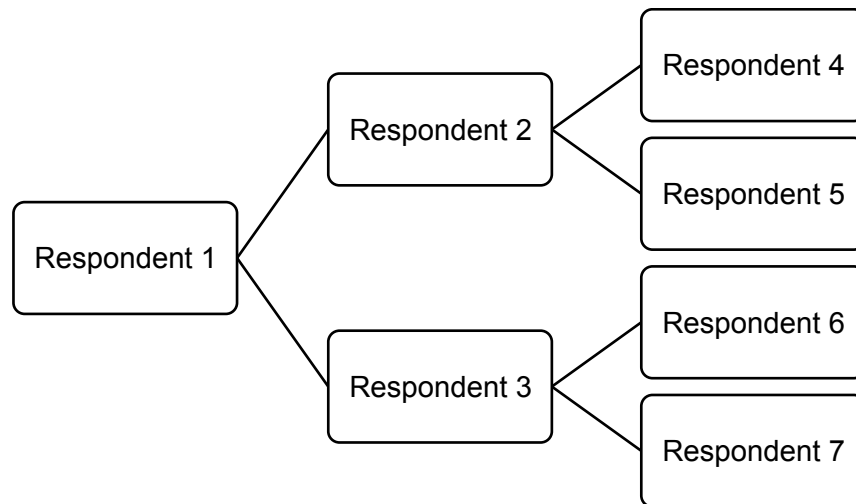


Figure 4. Exponential non-discriminative snowball sampling (Castillo, 2009).

In stratified purposeful sampling, also a non-random sampling strategy, the researcher selects information-rich cases for in-depth study (Patton, 2001). Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of the research. This sampling strategy allows the researcher to help illustrate characteristics of particular subgroups of interest and facilitates comparison between different groups. This method was adopted to ensure that I achieved a reasonable balance between the independent variables (e.g., male/female, tenured/contract) in the initial data gathering stage, which is an important assumption for the statistical procedures used in this study.

Obtaining a true representative random sample of the Japanese university English teaching population would be extremely difficult; therefore, I began gathering data by approaching nearly 50 friends, professors, acquaintances, and colleagues working at various Japanese universities and asked them to complete the questionnaire. I then asked them to either (a) distribute a paper version of the questionnaire or (b) to provide the link to the online version of the questionnaire to as many other teachers that they know or have contact with that fit the criteria as possible. This method of collecting data enabled me to gather roughly half of the total number of respondents. As a member of the Japanese Association of College English Teachers (JACET), I was given access to the membership directory and sent over 4,000 e-mail requests (with the online survey link attached) to the 2011 JACET members to collect the remaining participants' data. I received a 5% response rate that enabled me to collect roughly 200 more responses increasing the total number of respondents to over 400, a goal I had set from the beginning of the study. Tables 2 through 5 detail the descriptive statistics and bio-data of

the original respondents pre-data screening.

Table 2. *The Participants' Native Language and Gender*

Native language	English	Japanese
Number of respondents	225	215
Male	145	106
Female	80	109

Table 3. *The Participants' Teaching Experience and Gender*

Years of experience	1-10	11-20	21+
Number of respondents	127	192	121
Male	74	94	83
Female	53	98	38

Table 4. *The Participants' Tenured Vs. Contract Status and Gender*

Job rank	Tenured	Contract
Number of respondents	178	262
Male	96	155
Female	82	107

Table 5. *The Participants' Gender*

Gender	Male	Female
Number of respondents	251	189

Interviews

In order to answer research questions 5 and 6, I conducted interviews with 12 university English teachers. The interviewees were not selected based on their level of self-efficacy but rather carefully chosen to ensure all categories were represented equally. The four categories (or independent variables) in this study are Japanese speakers and

English speakers, experienced and inexperienced teachers, tenured and contract teachers, and male and female teachers. The justification for choosing 12 people to interview is that every individual fills multiple roles. For example, the same person might be female, tenured, Japanese, and have many years of teaching experience. The interviews were strictly optional and of the 440 teachers who responded to the survey, only 15% indicated that they would be willing to be interviewed. After initially scanning the list of 60 to 70 willing interview participants and dividing them into groups to ensure they covered all four categories (see above), 30 participants in geographically accessible areas to the researcher were contacted. From those 30 teachers, 12 teachers representing the four groups agreed and were available to be interviewed. Other factors affecting the final list of 12 interviewees included time restrictions on the part of the teachers and interviewer, availability, and physical location. In the end, all participants taught at institutions in either the Kansai or Kanto regions of Japan.

All data collection was conducted over a six-month period, from September 2011 through February 2012. A total of 440 questionnaire responses were gathered from teachers from 157 Japanese universities in the quantitative phase, and 12 university English teachers were interviewed in the qualitative phase. The online version of the questionnaire was put on Survey Monkey™, an online survey company that ensures that there can be no missing responses, increases the ease of the snowball sampling method, and removes the need for manual data input, significantly reducing potential data entry errors. Finally, although the ‘online consent form’ included in the online version of the questionnaire was not technically a consent form, it did however draw the participants’ attention to the fact that their names, institutions, and other personal information were not

necessary in order to complete the survey. Furthermore it outlined the precautions that I used in order to protect the data.

Ethics

All interview participants in this study were ensured of their anonymity and that participation in this project was completely voluntary. They understood that they had the right to withdraw and discontinue participation at any time without penalty or need for explanation. Furthermore, it was made clear to them that should they decline to participate or withdraw from the study, none of their data would be used. All participants understood that if they felt uncomfortable in any way during an interview session, they had the right to decline to answer any question or to end the interview. They also knew that the interviews would be digitally recorded and that the recordings would be kept secure and locked in the researchers desk until the completion of the project, at which time the recordings would be erased. Finally, the participants were never asked to state their names or institutions on the recording, and were ensured that their real names or institution names would never be used or identified in any reports using information obtained from this study.

Instrumentation

I developed a questionnaire based on the well established and validated Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Woolfolk Hoy, 2001). The 24-item TSES, measuring student engagement, instructional strategies, and classroom management, was modified to fit the English teaching situation at Japanese universities

(see Chapter 2 for a detailed description on the development of the TSES). In some cases, items needed nothing more than slight rewording or adapting; however, in other cases, the deletion of irrelevant (i.e., not applicable to university teachers) existing items and their replacement with new items was required. The benefit of adding new original items to the scale allowed me to ‘personalize’ the instrument so that it was well targeted for the university English teacher population in Japan.

Additionally and well before the creation and piloting of the instrument, following Bandura’s (1997) advice on how to measure teaching self-efficacy, an open-ended ‘pre-pilot’ survey investigating what impediments prevented teachers from doing their jobs to the best of their abilities was administered. The survey was given to approximately 50 university English teachers in Japan in late 2009, half of which were Japanese and half who were native English speakers. The teachers varied in age and experience. Aside from mentioning such obstacles as problem students, lack of motivation, and difficult language assessments (i.e., all of which are already included in the well established TSES), one of the most common and noticeable responses to the questionnaire involved issues with teachers’ superiors and the administrations at their respective institutions. Due to the majority of teachers citing their superiors and administrations as impediments, it was deemed necessary to include such a construct in the questionnaire. Thus, eight original items concerning administrative matters at Japanese universities were created, piloted, and added to the survey. The new 32-item scale, now measuring four constructs, was named the JULTEBS (Japanese University Language Teacher Efficacy Beliefs Scale) (see Table 6 or Appendix E). The respondents were asked to rate the 32 statements on a 6-point agreeability scale: 1 = *Strongly disagree*; 2 = *Disagree*; 3 = *Slightly disagree*; 4 =

Slightly agree; 5 = *Agree*; and 6 = *Strongly agree*. Following Bandura’s (2006) advice, teacher efficacy is measured against levels of task demands that represent gradations of challenges or impediments to successful performance. Consistent with self-efficacy’s theory of perceived competence, all items are written in present tense as ‘can do’ statements. Teachers were asked to judge their operative capabilities as of now, not their potential capabilities or their expected future capabilities. Furthermore, the self-efficacy appraisals reflect the level of difficulty teachers believe they can surmount. Finally, a 21-item background questionnaire (see Appendix B) was also administered, to collect such information as the respondents’ age, native language, gender, teaching experiences, and job status. The following is a concise visual timeline of the creation and administration of the surveys and interviews.

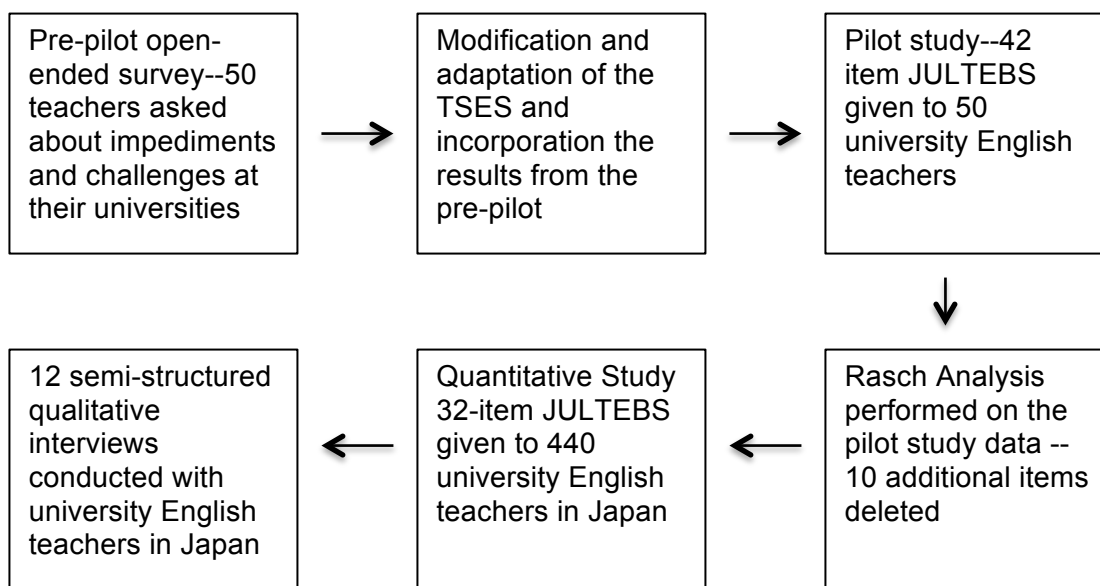


Figure 5. Timeline of the creation and administration of the JULTEBS.

Table 6. *Japanese University Language Teacher Efficacy Beliefs Scale (JULTEBS) Arranged by Construct (The Initial Version)*

Factors	JULTEBS items
Student Engagement	<ul style="list-style-type: none"> *1. I can communicate in a meaningful way with my students. +2. I can help my students think critically in English. +3. I can motivate most of my students to become more interested in English. *4. I can help students become more independent English learners. *5. I can get my students to believe they can succeed in English class. +6. I can get most of my students to value learning English. +7. I can foster student creativity. +8. I can greatly improve the overall English skills of my students.
Instructional Strategies	<ul style="list-style-type: none"> +1. I can respond effectively to English language questions from my students. +2. I can gauge my students' comprehension of what I have just taught. +3. I can provide appropriate support for my least proficient English language students. +4. I can use a variety of language assessment strategies in my classes. +5. I can provide appropriate challenges for my most proficient English language students. +6. I can put together level appropriate learning activities for my students. +7. I can adjust my lessons to different levels for individual students. +8. I can provide alternative explanations when students are confused.
Classroom Management	<ul style="list-style-type: none"> +1. I can control disruptive behavior in my English classroom. +2. I can make my expectations clear about appropriate student behavior. *3. I can create a learning friendly environment in my classroom. +4. I can deal effectively with defiant students. +5. I can establish effective routines in my classes. *6. I can get my students to follow classroom rules. +7. I can establish my own classroom management system. +8. I can prevent problem students from ruining classes.
Dealing with Superiors	<ul style="list-style-type: none"> *1. I can communicate my opinions effectively to my superiors. *2. I can convince my superiors to listen to my suggestions. *3. I can convince my superiors to act on my suggestions. *4. I can avoid being adversely affected by school bureaucracy. *5. I can convince my superiors to take my suggestions seriously. *6. I can get my superiors to support decisions I make regarding English teaching. *7. I can persuade my superiors to accommodate my teaching needs. *8. I can convince my superiors to have greater vision regarding language teaching.

Note. '*' = Original items (i.e., developed by the author and not taken from the TSES); '+' = Items taken from the TSES and modified.

Procedures

Most of the final 32 items in the JULTEBS were either modified from the TSES or created anew (Table 6). Therefore, it was necessary to establish the validity of the

instrument. As this instrument was directed at university language teachers, in September 2010, I individually asked four experienced university teachers in Japan to review a set of 50 items I had either adapted or created. The 50 items covered all four hypothesized constructs with at least ten items designed to measure each construct. I asked the teachers to identify problematic areas that were ambiguous, inappropriate, redundant, incomprehensible, or had missing information, as well as for overall advice on how the items could be improved. Based on the suggestions of the four teachers, eight of the 50 items were completely removed and nearly half of the remaining items were revised to various degrees. This left 42 items (Appendix D).

Having revised the questionnaire according to the teachers' suggestions, in November 2010, I piloted the 42 items with 50 university English teachers in Japan. After collecting the data and running a Rasch-based preliminary analysis (see Chapter 4 for a summary of the pilot study), ten additional items were deleted because they misfit the Rasch model, leaving 32 items or eight items per construct (see Appendix E).

Interviews

The qualitative data were gathered to help contextualize and explain the quantitative data and therefore act as a "follow-up" step to the questionnaire. Once the questionnaire data were collected, I conducted semi-structured, open-ended interviews with 12 teachers using a list of guiding questions (See Appendix H). The interviewees represented a sample of teachers with varying degrees of teacher self-efficacy. However, as the interviews were an optional part of the study and on a strictly volunteer basis, I was unable to secure interviews with teachers who had particularly high or low self-efficacy

scores. Instead, I interviewed both men and women, Japanese and foreign nationals, teachers who are tenured and those who had limited-term contracts, and teachers with varying degrees of university experience.

The interviews were conducted in English, digitally recorded with an Olympus Voice Trek DS-800, and lasted between 45 and 90 minutes. They were conducted like an informal conversation about issues pertaining to the English teaching situation at Japanese universities. The respondents were assured that their names and institutions would remain anonymous (see Appendix G for the interview consent form). Most interviews were held in the teachers' offices although several interviews were conducted in quiet cafés. I took brief notes as the participants talked and transcribed the audio recordings soon after completing the interview. I wrote what the interviewer and interviewee said as accurately as possible, but did not transcribe fillers, back channeling, false starts, or pauses, unless they were relevant or significant. I asked the teachers about their professional history and contextual factors at their institutions.

The goal of the interviews was to ask questions that elicit responses that help answer research questions 5 and 6 concerning support (or lack of support) that can build or challenge a teacher's self-efficacy. As the interviews were semi-structured and open-ended, care was taken to create a list of guiding questions (See Appendix H). Seventeen guiding interview questions were written. Questions 1 through 11, mainly inquired about impediments, hurdles, support, and backing in the participant's current teaching situation. Several of these questions were adapted from Borg (1998) and Law (1995), and several were created by me based on Bandura's (2006) advice that,

...assessing self-efficacy requires identifying the forms that challenges and impediments take...People are asked in open-ended interviews to describe the things that make it hard for them to perform the required activities regularly (p. 311)

Questions directly aimed at examining the sources of self-efficacy were also included. They can be found in interview questions 12 through 17. These questions were adapted from Mills' (2011) study, in which she examined major sources contributing to graduate French teaching assistants' self-efficacy in teaching literature. The interview questions were not intended to be asked mechanically from 1 to 17, nor was any pre-set order decided in advance.

Once the questions were created, they were sent to three experienced university EFL teachers in Japan with the instructions to answer the questions as best they could in about a one-hour sitting. Beside answering the questionnaire, the three teachers were asked to: (a) identify any problematic or ambiguous areas, (b) consider the likelihood that the questions would elicit the kind of information sought in research questions 5 and 6, and (c) give overall advice on how the questions could be improved. Based on their advice and suggestions, the guiding interview questions were directly related to the construct of self-efficacy as well as the themes included in the quantitative questionnaire.

Analyses

I first explain the three types of quantitative analyses used in this study (Rasch analysis, confirmatory factor analysis, and profile analysis) and discuss each method's respective criterion and benchmarks, chosen to evaluate the data. Next, I discuss the qualitative analysis and the method by which the interview data are to be examined.

Quantitative Analyses

The first step in confirming the validity and reliability of the questionnaire used in this study involved using the Rasch rating scale model (Andrich, 1978) to conduct an analysis of item fit and a Rasch Principal Component Analysis (PCA) of item residuals. Second, to fortify and to compare the results obtained from the Rasch analysis with a different type of analysis, I perform a Confirmatory Factor Analysis (CFA) to further check the dimensionality and reliability of the instrument. Lastly, in order to answer the first four research questions, I conducted a profile analysis on the data obtained from the questionnaire.

Rasch analysis.

The questionnaire data are analyzed with the Rasch rating scale model (Andrich, 1978). The formula for the Rasch rating scale model is as follows (Linacre, 2006, p. 13):

$$\log \{ P_{nij} / P_{ni(j-1)} \} = B_n - D_i - F_j, \text{ where}$$

\log is a natural logarithm, P_{nij} is the probability of respondent n scoring in category j for item i , $P_{ni(j-1)}$ is the probability of scoring in category $(j-1)$, B_n is the person measure of respondent n , D_i is the difficulty of item i , and F_j is the difficulty of category step j (the threshold at which there is a 50-50 chance of scoring in category j and category $j - 1$). The person's likely score is defined by the interaction between the person's measure, the item's difficulty, and the score's category threshold. Rasch analysis places persons (B_n) and items (D_i) on the same measurement scale where the unit of measurement is the logit (logarithm of odds unit).

Rasch person reliability is an estimate of the replicability of person placement that can be expected if the same respondents were to be given another set of items measuring the same construct. Rasch item reliability, on the other hand, is an estimate of the replicability of item placement within a hierarchy of items along the measured variable if these same items were to be given to another sample of comparable ability (Bond & Fox, 2007). Both reliability indices are analogous to Cronbach's alpha, and they range from 0 to 1. Generally, person and item reliability figures ranging from .91 to .94 are considered good, while reliability measurements greater than .94 are considered excellent (Fisher, 2007).

The Rasch person separation index is an estimate of the spread or separation of persons on the measured variable, and the item separation index is an estimate of the spread or separation of items on the measured variable. These indices provide a more sensitive measure of reliability because they are not bound by 0.0 and 1.0 like conventional reliability estimates (Bond & Fox, 2007). Higher values are considered better and a desirable minimum value for item separation is 2.0, as this indicates that item difficulties form at least two statistically distinct groups.

Item fit statistics are used to detect the extent to which the items match the predictions made by the Rasch model; items that fit the model well imply, but do not guarantee, the unidimensionality of the measured variable (Bond & Fox, 2007). Two Rasch fit statistics are commonly used: infit and outfit mean-square statistics. The item infit mean-square statistic is sensitive to the unexpected behavior of persons whose ability is at or near the item's difficulty estimate, and the item outfit mean-square statistic is sensitive to the responses of persons far above or below the item's difficulty. Linacre

(2009) describes good item model fit ranging from .5-1.5, and a very good item model fit falling between .7 and 1.4. This study uses the .5-1.5 range described by Linacre as a minimum item model fit criterion for both the infit and outfit mean-square statistics.

In addition to the mean-square fit statistics, Winsteps also provides standardized infit and outfit statistics. Unlike mean-square fit statistics, standardized fit statistics take into account *N*-size, and can have positive values indicating greater variation than suggested by the Rasch model, or negative values indicating less variation than expected. The ideal value is 0 with a standard deviation close to 1. The acceptable ranges for standardized infit and outfit statistics are greater than -2.0 and less than 2.0 (Bond & Fox, 2007). The standardized fit statistics will not, however, be used because they become too sensitive with large sample sizes such as the one in this study.

An item's goodness of fit to the Rasch model is one method of investigating the dimensionality of an instrument. However, a more effective approach to assessing the dimensionality of a set of items is through the use of a Rasch PCA of item residuals, as this approach identifies common variance among the items as well as the relationships among the residuals that remain after accounting for the primary component represented by the Rasch measures (Bond & Fox, 2007). It is important to remember not to interpret a Rasch item residual-based PCA as a usual factor analysis. Instead, these components show contrasts between opposing factors, not loadings on one factor (Linacre, 2009). In other words, a PCA of item residuals explains contrasting sub-structures in the data by breaking down the residual variance (Wright, 2000). If the variance explained by the Rasch measure is above 50% and the unexplained variance from the first contrast is less

than 5% or less than an eigenvalue of 3.0, the construct is considered fundamentally unidimensional (Linacre, 2009).

The Rasch model also makes it possible to produce a Wright map, which shows the items measuring each construct and the participants on a single interval logit scale. Wright maps are generated for each construct to provide a visual representation of the location of persons and items on the construct and to view the empirical item hierarchy. The empirical item hierarchies shown on the Wright map and the degree to which item difficulty estimates fit the participant ability estimates are discussed for each construct.

Finally, a Rasch Likert scale category functioning analysis is employed to determine whether the 6-point Likert scale employed in this study performed effectively. A 6-point Likert scale, with 1 representing *Strongly disagree* and 6 representing *Strongly agree*, was used with the both the pilot questionnaire and the main questionnaire in this study. The following criteria (Linacre, 2002) for good rating scale functioning were checked:

1. There are at least 10 observations for each step of the scale.
2. The average person measure for each step should be higher than the average person measure of the previous step.
3. Outfit mean square of each step should be less than 2.0.
4. There should be gaps in step difficulties of no less than .59 logits for a 6-point scale, .81 logits for a 5-point scale, and 1.1 logits for a 4-point scale.
5. Gaps in step difficulties should be less than 5.00 logits.

In the event the criteria for the 6-point scale were not met, Likert scale categories were collapsed until they met the criteria proposed by Linacre.

In sum, Table 7 (below) describes the important conditions for unidimensionality.

They are (a) item reliability and item separation are sufficiently high. Generally, values above .90 and 2.0 are regarded as providing acceptable reliability and separation, respectively. (b) No items misfit the Rasch model using the .5-1.5 infit mean-square criterion. (c) The variance explained by the measures is sufficiently high (above 50%) and the unexplained variance explained by the first contrast is less than 5% or less than an eigenvalue of 3.0 (Linacre, 2009). Using these criteria, I checked the dimensionality of each hypothesized construct in the questionnaire.

Table 7. *Rating Scale Instrument Quality Criteria*

Criteria	Critical value
Item measurement reliability	.91 to .94 is very good; > .94 is excellent
Item strata separation	Minimum of 2.00; higher numbers are better
Item model fit mean-square range	.5 to 1.5 is acceptable; .7 to 1.4 is very good
Variance explained by the Rasch measures	> 50% is good
Unexplained variance explained by first contrast (eigenvalue size)	< 3.0 is good

Confirmatory factor analysis.

A confirmatory factor analysis (CFA) is a theory-testing as opposed to a theory-generating method, such as exploratory factor analysis. In CFA, the researcher begins with a hypothesis prior to the analysis. This hypothesis specifies which variables are correlated with which factors and which factors are correlated. The hypothesis is based on a strong theoretical and/or empirical foundation (Stevens, 1996).

In addition, CFA offers researchers a more viable method for evaluating construct validity, as they are able to explicitly test hypotheses concerning the factor structure of

the data due to having the predetermined model specifying the number and composition of the factors.

Confirmatory methods, which are applied after specifying the factors beforehand, aim to match the observed and theoretical factor structures for a given data set in order to determine the “goodness of fit” of the predetermined factor model. Gorsuch (1983) explained how confirmatory factor analysis is powerful because it provides explicit hypothesis testing for factor analytic problems. Confirmatory factor analysis is the more important of the two major factor analytic approaches and should be the much more widely used. He specified that exploratory methods should be reserved only for those areas that are truly exploratory, that is, areas where no prior analyses have been conducted.

A CFA relies on several statistical tests to determine the adequacy of model fit to the data. It is important to remember when interpreting the findings from a confirmatory factor analysis that more than one model can adequately fit the data. Therefore, finding a model with good fit does not necessarily mean that the model is the only, or optimal model for that data. Therefore, because none of the fit indices are perfect and relying on only one would be unwise, fit should be simultaneously evaluated from the perspective of multiple fit statistics (Campbell, Gillaspay, & Thompson, 1995). The various fit indices’ strengths and weaknesses differ, therefore, if researchers can satisfy multiple approaches to estimating fit, then there is a higher probability they have identified a plausible model. Similarly, the use of several indicators follows Bollen’s (1989) and Joreskog’s (1993) recommendation of examining the extent to which the pattern of indicators is supportive of the model rather than relying on a single indicator of fit. In this study, the following

three indicators are used in judging the goodness of fit of the model: Chi-square (χ^2), the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA).

Chi-square (χ^2) tests the hypothesis that the model is consistent with the pattern of covariation among the observed variables. In the case of the chi-square statistic, smaller rather than larger values indicate a good fit. The chi-square statistic is sensitive to sample size, making it unclear in many situations whether the statistical significance of the chi-square statistic is due to poor fit of the model or to the size of the sample. This uncertainty has led to the development of many other statistics to assess overall model fit (Stevens, 1996). The Chi-square results are presented in this study in order to be comprehensive.

The Comparative Fit Index (CFI) compares how much better the model fits compared to a baseline model, typically the null model in which the observed variables are assumed to be uncorrelated (Kline, 2005). CFI ranges from 0 to 1 with a larger value indicating better model fit. Simulation studies by Hu and Bentler (1999) suggested that values above .90 indicate good fit of a model to the data. Thus, CFI values above .90 indicate adequate fit in this study.

The third statistic to measure overall fit in this study is the Root Mean Square Error of Approximation (RMSEA). The RMSEA takes into account the error of approximation in the population, is not affected by sample size, and is a measure of discrepancy per degree of freedom (Byrne, 2006). RMSEA values range from 0 to 1 and according to Hu & Bentler (1999) should ideally be .06 or below. However, Byrne (2006) indicated that an upper bound value of .08 for RMSEA could indicate adequate fit.

Finally, the 90% confidence interval for the RMSEA indicates the precision of the measurement and should be as small as possible. For this study, RMSEA indices below .08 were considered to show acceptable fit.

Profile analysis.

After the Rasch and confirmatory factor analyses were used for validating and refining the Japanese University Language Teacher Efficacy Beliefs Scale, the English teachers' responses to the questionnaire are analyzed using a profile analysis. The principal question for profile analysis is whether groups have different profiles on a set of measures. Profile analysis is a special application of multivariate analysis of variance (MANOVA) to a situation where there are several dependent variables, all measured on the same scale (Tabachnick & Fidell, 2007). In this study, the dependent variables are the four self-efficacy sub-constructs measured on the JULTEBS, while the independent variables are the four groupings of teachers: (a) native English teachers and non-native English teachers, (b) experienced and non-experienced teachers, (c) tenured and non-tenured teachers, and (d) male and female teachers.

All measures in a profile analysis must have the same range of possible scores, with the same score value having the same meaning on all of the measures (Tabachnick & Fidell, 2007). In this study, all four variables of teachers' self-efficacy use the same Rasch person ability estimates. In order to standardize the Rasch measures, the following steps were taken to anchor the rating scale and anchor the difficulty estimates on the same logit scale.

1. Analyze all items together. All items measuring the four variables are put into the

same analysis. These are the items I will have already validated, so I know they fit the Rasch model and form reasonably unidimensional constructs.

2. Obtain Likert category anchor values. Look at the category structure information for this analysis and see if the categories are working properly. Collapse categories if needed. When the categories look OK, record the Rasch-Andrich thresholds. These numbers are needed later.
3. Obtain and record the item difficulty estimates for all the items. These numbers are also needed later.
4. Work with the first construct, Student Engagement, by using the “SE only” command file. Use the category values from step 2 to anchor the Likert scale (I want to use the same Likert scaling for all four of my variables).
5. Add the item difficulty estimates for the SE items from step 3 to my SE command file. These are the item difficulty anchor values. Add the appropriate command to the command file. (The item difficulty estimates are the estimates from step 3).
6. When the command file is run, a column in the item statistics output says “Displacement” is now visible. Any values over .50 mean that the item should not be anchored (i.e., the item should be taken out of the IAFILE command).
7. Run this procedure for the remaining three variables.

In sum, by anchoring the rating scale and anchoring the item difficulty estimates on the same logit scale, the person ability estimates are permitted to “float,” but everything is on one logit scale, so the person ability estimates presented can be used as is in the profile analysis.

A profile analysis answers three types of questions. The first question is the flatness hypothesis. This concerns the similarities of responses and whether or not all the dependent variables elicit the same average response. In the context of this study, for example, the flatness test could evaluate whether male teachers and female teachers have the same pattern of self-efficacy scores, and whether or not male teachers as a whole, score notably high or low on any of the JULTEBS sub-constructs. The second question and arguably the most important, is the test of parallelism. This asks if different groups have parallel profiles. In the context of this study, for example, this test indicates if tenured teachers and contract teachers have the same pattern of highs and lows on the self-efficacy scale. The final question posed by a profile analysis is the levels hypothesis. This concerns whether or not one group, on average, scores higher on the selected set of measures than another. In the context of this study, for example, this can indicate if more experienced teachers have reliably higher self-efficacy scores than less experienced teachers.

Finally, the assumptions made in profile analysis are similar to those made when using MANOVA. Prior to conducting a profile analysis, Tabachnick and Fidell (2007) recommend the following assumptions be checked.

1. Sample size: There should be more research units in the smallest group than there are dependent variables.
2. Multivariate normality: Unless there are fewer cases than dependent variables in the smallest group and highly unequal n, deviations from normality of sampling distributions is not expected.

3. Absence of outliers: Profile analysis is extremely sensitive to outliers. Therefore, tests for univariate and multivariate outliers must be applied to all dependent variables.
4. Homogeneity of variance-covariance matrices: Unless sample sizes are highly unequal or there is evidence of strong homogeneity of the dependent variables, this assumption is probably safely ignored.
5. Linearity: For the relationships among the dependent variables, linearity is assumed. Evaluate this assumption by examining scatterplots between all pairs of dependent variables.
6. Absence of multicollinearity and singularity: Correlations among dependent variables are expected to be quite high when they are the same measure taken from the same cases over time. Only statistical multicollinearity poses difficulty.

Qualitative Analysis

The data were collected through open-ended interviews for the purpose of obtaining the teachers' insights and additional information about their self-efficacy. During the process of data analysis, the interview data recordings were listened to and transcribed verbatim. There were approximately 18 hours of recordings. To ensure accuracy, the transcriptions were proofread while listening and in the case of any confusion or misunderstandings, participants were contacted for confirmation and clarification of their responses. Patterns and themes found in the participants' interviews were identified through inductive coding methods explained by Tesch (1990). One advantage of using such a coding method is that the step-by-step approach engages a researcher in a systematic process of analyzing textual data. Each interview can be

reviewed line by line, and categories are generated using the seven-step coding process described below (Tesch, 1990, pp. 142-145).

1. Get a sense of the whole. Read all transcriptions carefully.
2. Choose the data from any one interview and as you go through ask yourself, “What is this about?” Try not to think about the substance of the information, but more on it’s underlying meaning.
3. Complete the second task again with the data from several more interviews and make a list of all topics. Cluster similar topics together and make a list of major/minor topics, unique topics, leftovers, etc.
4. Take the list back to your data and abbreviate the topics you’ve created as codes. Write the codes next to the appropriate segments of the interview transcriptions. Be aware of any new categories that emerge.
5. Find the most descriptive wording for your topics and turn them into categories. Look for ways to reduce your total list of categories by grouping topics that relate to each other.
6. Make a final decision on the abbreviation for each category and alphabetize these codes.
7. Assemble the material belonging to each category in one place.

Following this coding process, the identified recurring themes and categories (e.g., those that were surprising or not anticipated at the beginning of the study) were investigated separately from the questionnaire data and later examined in conjunction with the

research questions of the present study for the purpose of refining, extending, and/or contextualizing the survey data.

Summary

In this chapter, I explained the type of mixed-methodology employed and described the participants, instrumentation, procedures, and analyses. In the next chapter, I report on the preliminary analyses and instrument validation of the quantitative data.

CHAPTER 4

INSTRUMENT VALIDATION

Three main steps were conducted during the instrument validation process. In the first step, the Rasch rating scale model (Andrich, 1978) was used to analyze data from the pilot study in order to select the strongest items for the Japanese University Language Teachers Efficacy Beliefs Scale used in the main study, as well as to confirm the existence of the four hypothesized factors. Additionally, for the sake of the instrument's balance and brevity, a final aim of the preliminary analysis was to trim any misfitting, redundant, or unnecessary items, so that the final number of items used in the questionnaire was as parsimonious as possible. These results are briefly summarized first. In the second step, the Rasch rating scale model was used to examine the validity and reliability of the items included in the main questionnaire and to reconfirm the presence of the four constructs. In the third step, subjecting the data to a Confirmatory Factor Analysis, and then comparing the results to those of the Rasch analysis further explored the validity of the main questionnaire.

Rasch Analysis of Item Fit and PCA of Item Residuals

To examine the construct validity of both the pilot study questionnaire and the revised main questionnaire, the Rasch measurement model was employed using WINSTEPS version 3.70.0.2 (Linacre & Wright, 2009). In the pilot study, due to the small *N*-size ($N = 50$) and item to person ratio (nearly 1 to 1), an ordinary factor analysis through SPSS was not a viable option, as it would not have met the assumptions required

to check the hypothesized factors. Because the Rasch measurement model's precision is not severely compromised by small sample sizes and it still produces accurate item fit and dimensionality statistics with 50 participants, it was chosen for use at this stage of the study.

Pilot Study Analysis

In the pilot study, 50 university English teachers in Japan, half of whom were native English speakers and half who were native Japanese speakers, completed a 42-item teacher self-efficacy questionnaire. A Rasch analysis of item fit and a Rasch principal component analysis (PCA) of item residuals was performed on the four hypothesized factors (Efficacy in Student Engagement (SE), Efficacy in Instructional Strategies (IS), Efficacy in Classroom Management (CM), and Efficacy in Dealing with Superiors (DS)) the questionnaire was designed to measure. The order in which the constructs were subjected to the Rasch analysis was determined by first checking the dimensionality of all 42 items. It was hypothesized that the 42 items would not be unidimensional and that this would be reflected in a large eigenvalue (i.e., > 3.0) in the first contrast. In this initial analysis, the Rasch model accounted for 50.5% of the variance (eigenvalue = 42.0) and the first residual contrast accounted for 10.3% of the variance (eigenvalue = 8.5). All of the items from one hypothesized construct that most strongly showed contrast from all other constructs were removed and analyzed first. Next, the dimensionality of the remaining items from the other three hypothesized constructs were checked. Again, all of the items from one construct that most strongly showed a contrast from all other constructs were removed and analyzed. The process was repeated

until all items had been removed and analyzed. The resulting order of analysis was: Dealing with Superiors, Student Engagement, Classroom Management, and Instructional Strategies. The following is a brief summary of the pilot study results.

Efficacy in Dealing with Superiors

Twelve items from the pilot survey were designed to measure the hypothesized factor, Efficacy in Dealing with Superiors. A Rasch analysis was conducted for the 12 items. Based on the results of the analysis, four items (DS 5, *I can choose which classes I want to teach* (Infit MNSQ = 2.01; Outfit MNSQ = 1.86); DS 8, *I can fail students without approval* (Infit MNSQ = 2.36; Outfit MNSQ = 1.92); DS 11, *I can get my superiors to accommodate my desired teaching schedule* (Infit MNSQ = 1.82; Outfit MNSQ = 1.76); and DS 4, *I can deal with rules created by my superiors* (Infit MNSQ = 1.65; Outfit MNSQ = 1.59) were deleted because they did not fall between the .5-1.5 infit mean square criterion range, outlined by Linacre (2009). The remaining eight items fell between the criterion range and they appeared to form a fundamentally unidimensional construct, as the variance explained by the Rasch model was to 67.1%, the unexplained variance in the first contrast was 9.8%, and the eigenvalue was below the 3.0 benchmark at 2.4. Item reliability and separation were high at .95 and 4.37, respectively.

Efficacy in Student Engagement

Eight items from the pilot survey were designed to measure the hypothesized factor, Efficacy in Student Engagement. A Rasch analysis was conducted for the eight items. Based on the results of the analysis, no items were deleted. However, the results

also showed slight item clustering due to some of the items being too easily endorsed. Therefore, the wording of three items (SE 3, SE 6, and SE 8) was changed slightly to make them more difficult to endorse (see Appendices D and E for the “before” and “after” wording of the three items). The original eight items measuring Efficacy in Student Engagement were considered fundamentally unidimensional because the amount of variance explained by the Rasch measures was 56.6%, which was above the 50% criterion. The unexplained variance in the first contrast was 11.3% and the eigenvalue was 2.1, which was below the 3.0 criterion. Rasch item reliability was a satisfactory at .89 and item separation was 2.89, which was above the 2.0 benchmark. Finally, the infit mean square statistic for all items fell between the .5-1.5 criterion range.

Efficacy in Classroom Management

Nine items from the pilot survey were designed to measure the hypothesized factor, Efficacy in Classroom Management. A Rasch analysis was conducted for the nine items. Based on the results of the analysis, item CM 7 (*I can deal effectively with uncooperative students*) was chosen for deletion because of its logical redundancy with item CM 4 (*I can deal effectively with defiant students*). Both items seemed to measure the same construct. After removing the one item, the remaining eight items measuring Efficacy in Classroom Management were considered fundamentally unidimensional because the variance explained by the Rasch model was to 58.5% with an eigenvalue of 11.3, the unexplained variance in the first contrast was 11% and the eigenvalue was below the 3.0 benchmark at 2.1. Rasch item reliability and separation were satisfactory at .87 and 2.55, respectively.

Efficacy in Instructional Strategies

Thirteen items from the pilot survey were designed to measure the hypothesized factor, Efficacy in Instructional Strategies. A Rasch analysis was conducted for the 13 items. Based on the results of the analysis, five items (IS 3, *I can keep activities running smoothly for the entire lesson*; IS 4, *I can teach an English language concept better after having taught it once before*; IS 5, *I can improvise when necessary in my English language classes*; IS 8, *I can gauge my students' English progress over the semester*; and IS 10, *I can implement alternative teaching strategies in my classroom to accommodate the various levels of students*), were eventually deleted because they did not fall between the infit mean square criterion ranges .5-1.5 proposed by Linacre (2009). After removing five items, the remaining eight items measuring Efficacy in Instructional Strategies fell between the acceptable criterion range and appeared to form a fundamentally unidimensional construct. The variance explained by the Rasch model was 61.4% (eigenvalue = 12.7), the unexplained variance in the first contrast declined to 9.4%, and the eigenvalue was 2.1, which is below the 3.0 benchmark. Item reliability and separation were both strong at .92 and 3.34, respectively.

Summary of Pilot Study Rasch Analysis

In conclusion, through the use of the Rasch model item statistics, the four hypothesized factors were confirmed to be fundamentally unidimensional. The results are summarized in Table 8.

Table 8. A Summary of the Four Factors Identified by the Pilot Study Questionnaire

Name	Label	Item number	Variance accounted for by the Rasch Model (%)	Eigenvalue of the first residual	IR	IS
Dealing with Superiors	DS	1, 2, 3, 6, 7, 9, 10, 12	67.1	2.4	.95	4.37
Student Engagement	SE	1, 2, 3*, 4, 5, 6*, 7, 8*	56.6	2.1	.89	2.89
Classroom Management	CM	1, 2, 3, 4, 5, 6, 8, 9	58.5	2.1	.87	2.55
Instructional Strategies	IS	1, 2, 6, 7, 9, 11, 12, 13	61.4	2.1	.92	3.34

Note. * = The wording of the item was modified; IR = Item reliability; IS = Item separation.

Data Screening

Data screening was completed before validating the main instrument through both a Rasch analysis and Confirmatory Factor Analysis. The assumptions for the CFA were checked and met. The following steps were taken prior to running the confirmatory factor analysis.

Applying a common rule of thumb for a factor analysis where the person to item ratio should be at least 10 to 1 (Velicer & Fava, 1998) the first assumption concerning sample size was met with a starting *N*-size of 440 and a 32-item questionnaire. Next, as the questionnaire was conducted primarily online using Survey Monkey™ and because I was able to control the settings so that all questions had to be completed before the participant could submit their responses, there were no missing data. Furthermore, by virtually eliminating the need for manual data input, potential data entry errors were significantly reduced.

Next, univariate outliers were checked using SPSS EXAMINE. Using Z-scores, five respondents had extreme values for all four constructs on the questionnaire, defined

here as cases that were more than 3.29 standard deviations above or below the mean (Tabachnick & Fidell, 2007). They were deleted from the data. Next, the data were checked for multivariate outliers using Mahalanobis distance in SPSS REGRESSION. With 31 degrees of freedom, the critical chi-square value at $p < .001$ is 61. After running five iterations, 37 respondents were eliminated as they had values greater than 61. Therefore, the total number of participants used for instrument validation was 398.

Instrument Validation: Rasch Analysis

The following Rasch statistics are provided in the following order for each construct: (a) rating scale functioning, (b) item fit statistics, (c) Principal Component Analysis (PCA) of item residuals, (d) item and person reliability and separation, and (e) the Wright map.

Construct 1: Efficacy in Dealing with Superiors

Eight items from the main survey measured the construct Efficacy in Dealing with Superiors (Appendix E). First, the Likert scale category functioning was examined (Table 9). The minimum of 10 observations per category was met, as the smallest number of observations was 91 (category 1). The outfit MNSQ statistic for all categories was well below the 2.0 criterion. Separation between adjacent thresholds was greater than the required .59 logits for a 6-point scale and there were no disordered thresholds. In sum, the six-category structure of the scale was appropriate and met the criteria set by Linacre (2002).

Table 9. *Category Structure Functioning for Efficacy in Dealing with Superiors*

	Count (%)	Infit MNSQ	Outfit MNSQ	Structure calibration	Category measure
1 Strongly disagree	91 (4)	.99	1.00	None	(-6.86)
2 Disagree	263 (11)	1.00	1.04	-5.73	-4.20
3 Slightly disagree	403 (17)	.84	.81	-2.60	-1.74
4 Slightly agree	895 (37)	.94	.95	-.89	.84
5 Agree	602 (25)	1.08	1.06	2.50	4.61
6 Strongly agree	134 (6)	1.27	1.13	6.72	(7.83)

Next, the infit mean square statistics and a PCA of item residuals were calculated for all DS items (see Table 10). Item DS 4 (*I can avoid being adversely affected by school bureaucracy*) displayed the worst fit (Infit MNSQ = 1.92; Outfit MNSQ = 1.89). Furthermore, not only was DS 4 misfitting, it was also deemed confusing based on its content. This item differed from the seven other items in two significant ways. First, the item does not use the word *superior* as used in all of the other items. Instead, it uses the more general term of *school bureaucracy*. Second, it differs from the other items because of its reverse coding. The positive words *convinces* or *persuades* were used with the other items, but the negative word *avoid* was used with this item. Subsequently, item DS 4 was deleted.

Table 10. *Rasch Item Statistics for Efficacy in Dealing with Superiors (pre-item deletion)*

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt-measure correlation
DS 8	1.11	.07	.74	-3.9	.76	-3.6	.84
DS 3	.84	.07	.62	-6.1	.64	-5.7	.85
DS 5	.28	.07	.51	-8.1	.53	-7.8	.89
DS 4	.04	.07	1.92	9.9	1.89	9.8	.63
DS 2	.00	.08	.79	-3.1	.80	-3.0	.84
DS 7	-.11	.08	.99	-.1	.98	-.2	.77
DS 6	-.15	.08	.84	-2.3	.81	-2.8	.82
DS 1	-2.01	.09	1.58	6.4	1.50	5.0	.62

After deleting item DS 4, the infit and outfit mean square statistics for the

remaining items were checked again. This time, item DS 1 (*I can communicate my opinions effectively to my superiors*) displayed the worst infit and outfit mean square statistic (Infit MNSQ = 1.89; Outfit MNSQ = 1.83). The item does make sense logically and when the Rasch person ability estimates were correlated with and without item DS 1, the item did not disturb the person measures, as the correlation was high at .99 ($p < .01$; 2-tailed test). The decision, however, was still made to delete the item based on its poor fit statistics. Furthermore, there were still more than enough (i.e., 6 items) items left to measure the construct. In total, two of the original eight items were deleted. The remaining six items (DS 2, DS 3, DS 5, DS 6, DS 7, DS 8) were subjected to the same analysis as above. All of the remaining items fit the Rasch model and the part-measure correlations ranged between .80 and .89 (Table 11).

Table 11. *Rasch Item Statistics for Efficacy in Dealing with Superiors (post-item deletion)*

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt-measure correlation
DS 8	1.17	.09	1.04	.6	1.08	1.0	.87
DS 3	.77	.09	.75	-3.6	.74	-3.7	.89
DS 5	-.08	.09	.69	-4.7	.68	-4.7	.91
DS 2	-.49	.09	1.01	.1	.97	-.3	.87
DS 7	-.65	.09	1.39	4.7	1.38	4.5	.80
DS 6	-.71	.09	1.05	.7	1.02	.3	.86

After deleting item DS1, the variance explained by the Rasch model increased to 75.9% (eigenvalue = 20.1) and the unexplained variance in the first contrast dropped to 7.6% and the eigenvalue was below the 3.0 benchmark at 1.9. Thus, the items appeared to form a fundamentally unidimensional construct.

The Wright map for the Dealing with Superiors construct (Figure 6) shows the logit scale for both persons and items on the left side. The scale ranges from -8 to 9, with

an item mean of 0. Participants are represented on the left side of the Wright map as number signs (#) and periods (.), with each # standing for four persons and periods standing for one person. The items measuring Dealing with Superiors are displayed on the right side of the vertical line. Persons are placed along the line according to their Rasch person ability estimates, and items are placed according to their endorsement difficulty estimate. A person at the mean (M) of person ability on the left side has a 50% chance of endorsing an item of the same difficulty level (M) on the right side of the vertical line (Bond & Fox, 2007).

Items located above the item mean are more difficult to endorse so that a person located at the person mean has less probability of endorsing the item. The mean person ability ($M = .88$, $SD = 2.97$) and mean item difficulty ($M = .00$, $SD = .72$) showed that the items on the whole were well targeted for the sample. The easiest item to endorse was item DS 6 (*I can get my superiors to support my decisions I make regarding English teaching*, difficulty measure = $-.71$). The most difficult item to endorse was item DS 8 (*I can convince my superiors to have greater vision regarding language teaching*, difficulty measure = 1.17). These results likely reflect the frustration or lack of confidence that many respondents felt concerning getting their superiors to be more open-minded regarding English teaching. Logically however, it seems natural that the respondents felt confident that they could get their superiors to support their own individual teaching styles and decisions they make in the classroom. In addition, the difficulty hierarchy of items DS 2 (*I can convince my superiors to listen to my suggestions*), which was easier to endorse than DS 5 (*I can convince my superiors to take my suggestions seriously*), which

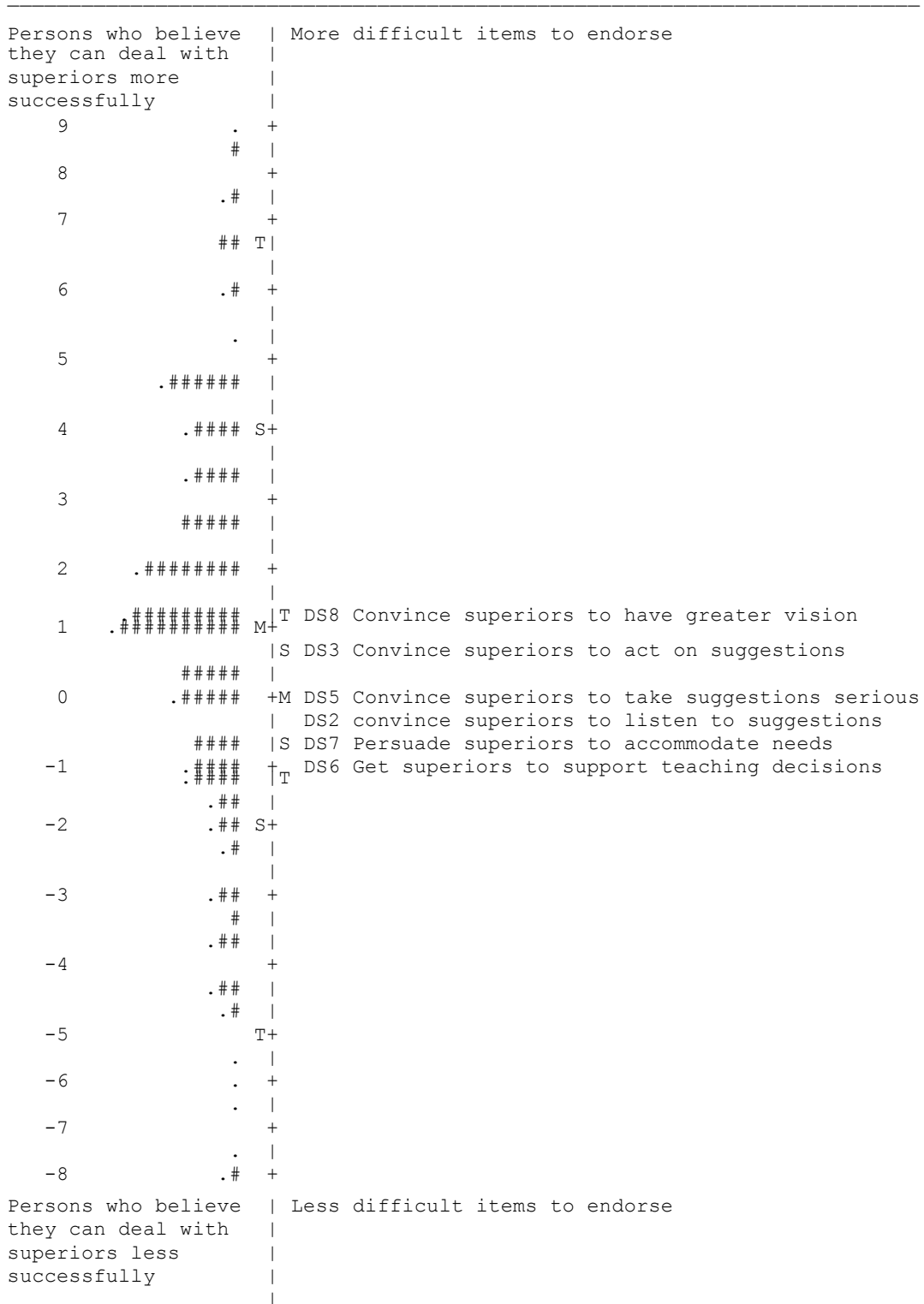


Figure 6. Wright map for Dealing with Superiors. Each “#” = 4 persons. Each “.” = 1 to 3 persons. M = Mean; S = 1 standard deviation; T = 2 standard deviations.

was easier to endorse than DS 3 (*I can convince my superiors to act on my suggestions*) makes sense because each one takes more effort to achieve than the next (i.e., listening is easier to do than taking something seriously, which is easier to do than acting on a suggestion). In sum, the Wright map shows the items are well centered on the persons and the item hierarchy is logical.

Finally, all items measuring this construct fell between the .5-1.50 criterion range of the infit and outfit mean square statistics, and item reliability (.98) and separation (7.68) were both high and above their respective benchmarks. The person reliability (.91) and separation (3.25) statistics were also above their benchmarks.

In sum, the six-category structure of the scale was appropriate and met the necessary criteria. Furthermore, after removing two items, the remaining six items measuring Efficacy in Dealing with Superiors were considered fundamentally unidimensional because the items' fit to the Rasch model was good, the variance accounted for by the Rasch model was high, the variance accounted for in the first residual contrast was below the 3.0 eigenvalue criterion, and item separation and reliability was strong.

Construct 2: Efficacy in Student Engagement

Eight items from the main survey were designed to measure Efficacy in Student Engagement. First, the Likert scale category functioning was examined. Although the outfit MNSQ statistic for all categories was below the 2.0 criterion and separation between adjacent thresholds was greater than the required .59 logits, the minimum of 10 observations per category was not met, as the smallest number of observations in

category 1 was 8. Therefore, category 1 (*Strongly disagree*) and category 2 (*Disagree*) were combined; the results can be seen in Table 12.

Table 12. *Category Structure Functioning for Student Engagement*

	Count (%)	Infit MNSQ	Outfit MNSQ	Structure calibration	Category measure
1 Disagree	86 (3)	1.04	1.06	None	(-4.49)
2 Slightly disagree	328 (10)	.92	.91	-3.25	-2.55
3 Slightly agree	1053 (33)	.93	.94	-1.79	.42
4 Agree	1238 (39)	.93	1.16	.87	2.53
5 Strongly agree	479 (15)	1.14	1.12	4.17	(5.30)

After combining the first two categories, the modified 5-point scale met the criterion based on Linacre’s guidelines, including the .81 logit criterion for separation between adjacent thresholds for a 5-point scale.

The infit mean square statistic (Table 13) for all items fell between the criterion range of .5 and 1.5, and the part-measure correlations ranged between .52 and .84.

Table 13. *Rasch Item Statistics for Efficacy in Student Engagement*

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt-measure correlation
SE 8	1.18	.08	1.02	.3	1.06	.9	.76
SE 2	.61	.08	1.25	3.2	1.25	3.3	.72
SE 3	.47	.08	.71	-4.4	.70	-4.6	.84
SE 6	.43	.08	.84	-2.3	.84	-2.4	.81
SE 4	.23	.08	.82	-2.5	.83	-2.4	.79
SE 7	-.04	.09	1.07	.9	1.06	.9	.74
SE 5	-.48	.09	.78	-3.1	.80	-2.9	.78
SE 1	-2.40	.10	1.40	4.9	1.83	6.1	.52

The amount of variance explained by the Rasch measures was 62.9% (eigenvalue = 13.5), which was above the 50% criterion. The unexplained variance in the first contrast was 7.4% and the eigenvalue was 1.6, which was below the 3.0 criterion. Thus, the construct was fundamentally unidimensional. Finally, the Rasch item reliability

estimate was .99 and item separation was 11.17 and well above the 2.0 benchmark. The Rasch person reliability estimate (.90) and separation (2.69) also satisfactorily met the required benchmarks.

The Wright map for the Dealing with Superiors construct is displayed in Figure 7. The scale ranges from -3 to 8, with an item mean of 0. A comparison of the locations of the means (M) for the person measures and item measures show that the mean person measure ($M = 2.24$; $SD = 1.97$) was much higher than the mean item measure ($M = .00$; $SD = 1.01$). This indicates that generally, the participants endorsed the items easily. In other words, they believed that they were highly capable of engaging their students in the situations described by the items. Item SE 1 (*I can communicate in a meaningful way with my students*; difficulty estimate = -2.40) was the easiest item to endorse while item SE 8 (*I can greatly improve the overall English skills of my students*; difficulty estimate = 1.18) was the most difficult item to endorse. As the respondents measured on this scale were entirely university teachers, it should not be surprising that item SE 1 was the easiest to endorse because most teachers at that level would not say or at least not want to admit that they cannot communicate meaningfully with their own students. Item SE 5 (*I can get my students to believe they can succeed in English class*), the second easiest item to endorse, is a very concrete idea (i.e., the least abstract). In other words, teachers who set appropriate and realistic goals in their English classes should be able to convince their students that they will be able to succeed, regardless of the students' ability. Similarly, item SE 6 (*I can get most of my students to value learning English*), which is slightly more difficult to endorse than item SE 5, is a less concrete idea because presenting

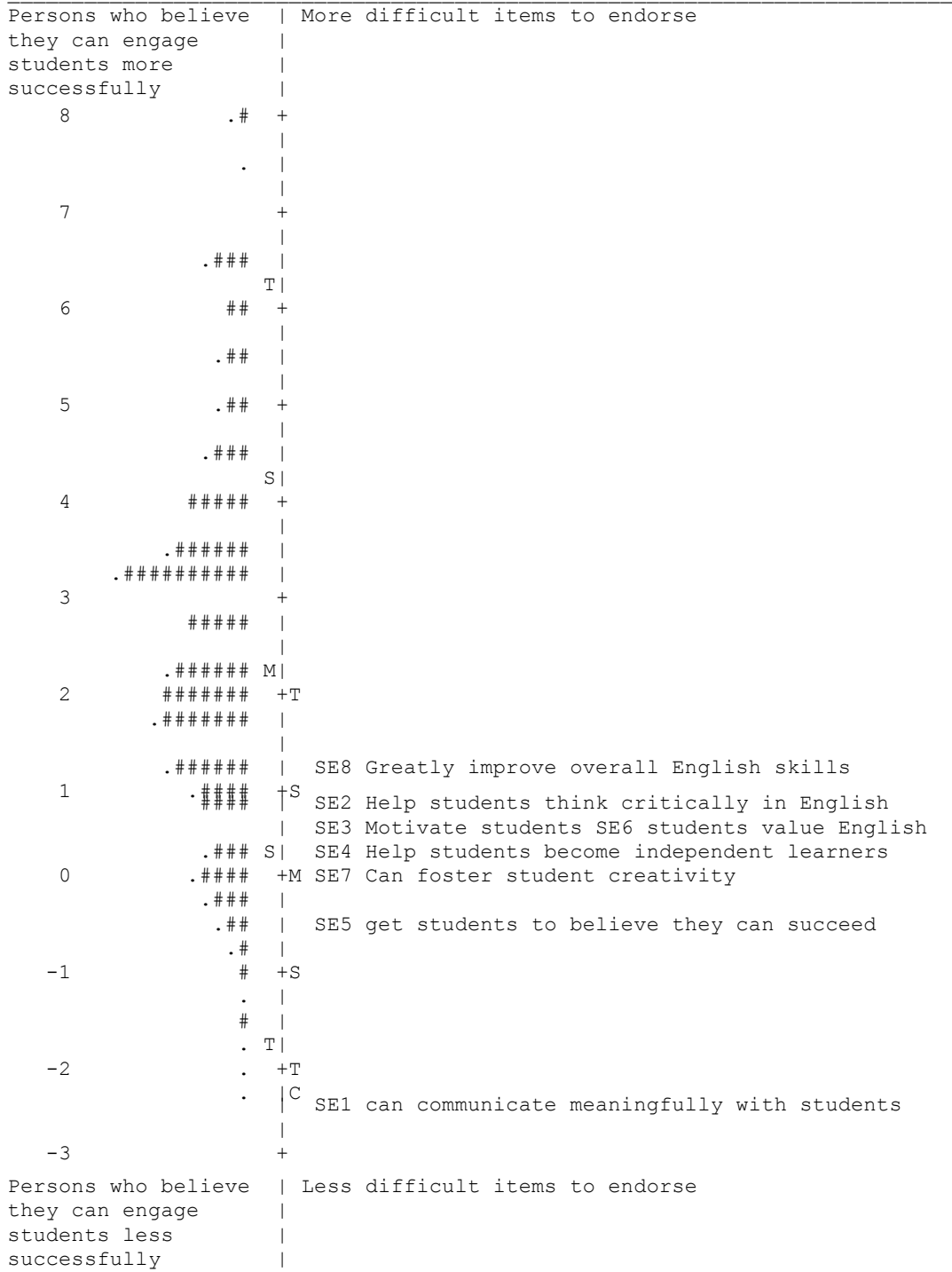


Figure 7. Wright map for Student Engagement. Each “#” = 4 persons. Each “.” = 1 to 3 persons. M = Mean; S = 1 standard deviation; T = 2 standard deviations.

students with the “bigger” picture and transmitting to them why English might have value to their lives might feel like a more difficult task. Furthermore, due to the even greater abstractness of item SE 2 (*I can help my students think critically in English*), it is logically more difficult to endorse than item SE 6. For example, a teacher might not be sure if their students are even capable of thinking critically in their native tongue, leaving the teacher unsure of how successfully they can get their students to think critically in English. Lastly, it is reasonable that item SE 8 is the most difficult to endorse because many teachers to some degree feel confident that they can improve some, but not all of their students’ skills.

In sum, the eight items measuring Efficacy in Student Engagement were considered fundamentally unidimensional because the items’ fit to the Rasch model was good, the variance accounted for by the Rasch model was above the 50% criterion measure, the unexplained variance accounted for in the first residual contrast was below the 3.0 criterion, and item separation and reliability were strong.

Construct 3: Efficacy in Classroom Management

Eight items from the main questionnaire were designed to measure Efficacy in Classroom Management. First, the Likert scale category functioning was examined. Although the outfit MNSQ statistic for all categories was below the 2.0 criterion and separation between adjacent thresholds was greater than the required .59 logits, the minimum of 10 observations per category was not met in two cases. The first case indicated that no respondent used the *Strongly disagree* category and the second case indicated that only four responses were recorded in the *Disagree* category. After

combining the first two categories, the minimum of 10 observations criterion was still not met. For this reason, category 1 (*Strongly disagree*) and category 2 (*Disagree*) were combined with category 3 (*Slightly disagree*) to form one general *Disagree* category. The modified category structure can be seen below in Table 14.

Table 14. *Category Structure Functioning for Classroom Management*

	Count (%)	Infit MNSQ	Outfit MNSQ	Structure calibration	Category measure
1 Disagree	88 (3)	1.02	1.04	None	(-4.56)
2 Slightly agree	561 (18)	1.02	1.03	-3.42	-1.90
3 Agree	1625 (51)	.95	1.00	.37	1.72
4 Strongly agree	910 (29)	.98	.97	3.79	(4.90)

After combining the first three categories, the resulting four-point scale met all of Linacre's criteria, including the 1.1 logit criterion for separation between adjacent thresholds for a 4-point scale.

The infit and outfit mean square statistics (Table 15) fell between the criterion range of .5-1.50, and the part-measure correlations ranged between .66 and .81

Table 15. *Rasch Item Statistics for Efficacy in Classroom Management*

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt-measure correlation
CM 4	1.85	.10	.90	-1.4	.92	-1.1	.81
CM 8	.68	.10	.94	-.8	.94	-.8	.77
CM 6	-.08	.11	.81	-2.8	.79	-2.8	.77
CM 1	-.14	.11	.96	-1.0	.92	-1.0	.78
CM 2	-.39	.11	.97	-.5	.96	-.5	.73
CM 5	-.50	.11	1.10	.7	1.06	.7	.71
CM 3	-.67	.11	1.30	4.4	1.43	4.4	.66
CM 7	-.74	.11	.98	.2	1.01	.2	.75

A PCA of item residuals was calculated for all CM items and the items appeared to form a fundamentally unidimensional construct. The amount of variance explained by

the Rasch measures was 57.2% (eigenvalue = 10.7), which was above the suggested value of 50%. The unexplained variance in the first contrast was 9.7% and the eigenvalue was 1.8, which was below the 3.0 criterion. Finally, the Rasch item reliability estimate was .98 and item separation was 7.45 and above the 2.0 benchmark. The Rasch person reliability estimate (.91) and separation (2.34) satisfactorily met their benchmarks as well.

The Wright map for the Classroom Management construct is displayed in Figure 8. The scale ranges from -2 to 8, with an item mean of 0. A comparison of the locations of the means (M) for the person ability measures and item difficulty measures show that the mean person ability measure ($M = 3.49$; $SD = 2.10$) was much higher than the mean item difficulty measure ($M = .00$; $SD = .82$), which indicates that the participants easily endorsed the items. This suggests that they believed that they were highly capable of managing their classrooms in the situations described by the items. Item CM 7 (*I can establish my own classroom management system*; difficulty estimate = $-.74$) was the easiest item to endorse, while item CM 4 (*I can deal effectively with defiant students*; difficulty estimate = 1.85) was the most difficult item to endorse. Logically, the item hierarchy in Figure 8 makes sense in that one would assume that university teachers would more likely be confident in creating a classroom management system (item CM 7: the easiest item) and creating a friendly environment (item CM 3: the second easiest item) than getting problem students to follow rules (item CM 6: the third most difficult item) or dealing with defiant students (item CM 4: the most difficult item). Furthermore, the four easiest-to-endorse items (CM 7, 3, 5, and 2) all deal with what the teacher does individually, whereas the four most difficult-to-endorse items (CM 4, 8, 6, 1) all involve

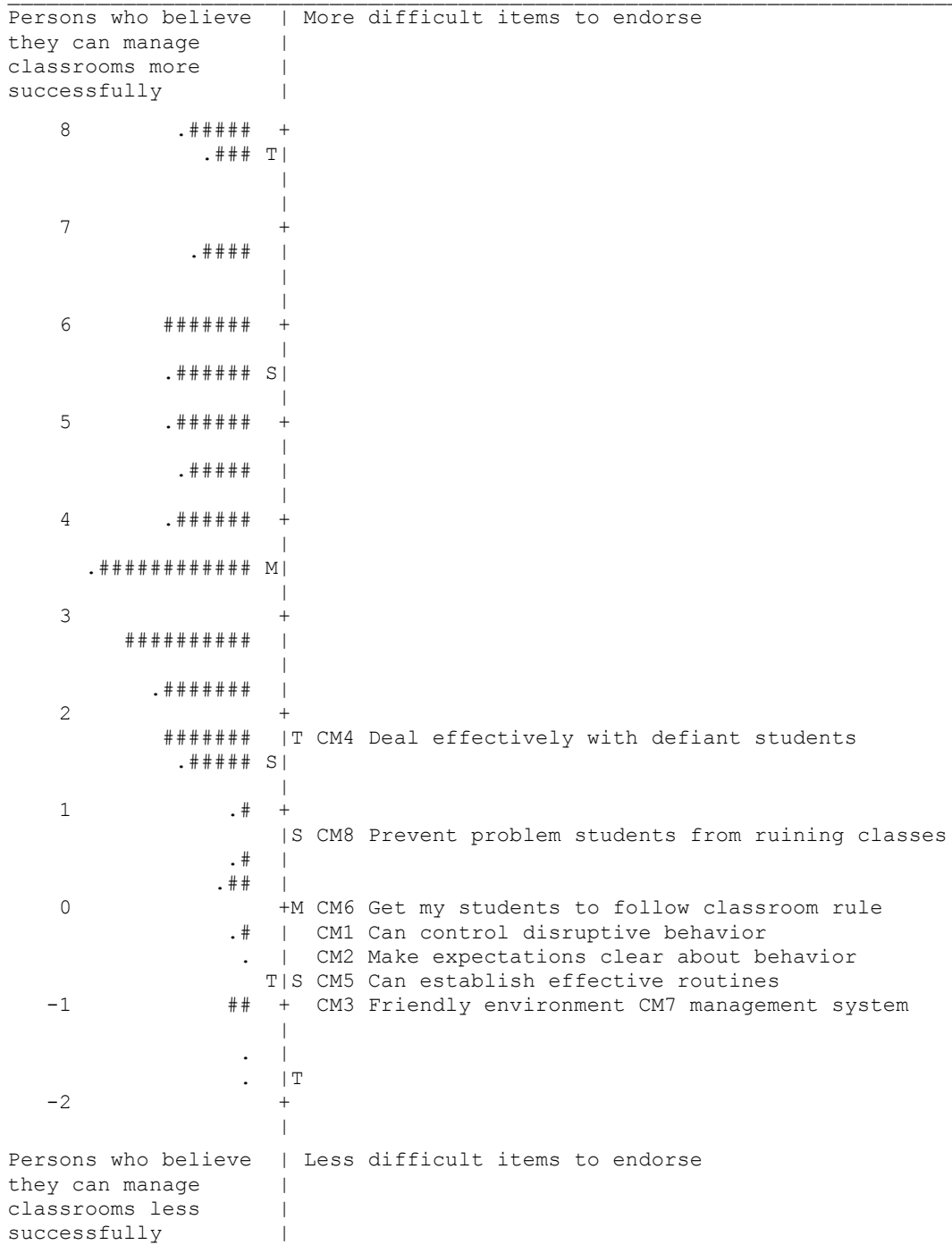


Figure 8. Wright map for Classroom Management. Each “#” = 4 persons. Each “.” = 1 to 3 persons. M = Mean; S = 1 standard deviation; T = 2 standard deviations.

dealing with students directly. Furthermore, the Wright map suggests a ceiling effect with the items; persons with high ability estimates (i.e., those who felt confident in their classroom management abilities) were not measured as precisely because not enough difficult-to-endorse items were available to measure the persons at that level of ability. The lack of difficulty in answering items for this construct might suggest that classroom management at the university level is not as major of an issue as it might be for example, at the secondary level.

In sum, although the items measuring Efficacy in Classroom Management were easily endorsed, they were considered fundamentally unidimensional because the items' fit to the Rasch model was good, the variance accounted for by the Rasch model was acceptable, the eigenvalue of the first residual contrast was below the 3.0 criterion, and item and person reliability and separation were strong.

Construct 4: Efficacy in Instructional Strategies

The final eight items from the main survey were designed to measure Efficacy in Instructional Strategies. First, the Likert scale category functioning was examined. Although the outfit MNSQ statistic for all categories was below the 2.0 criterion and separation between adjacent thresholds was greater than the required .59 logits, the minimum of 10 observations per category was not met. There were no observed responses in category 1 (*Strongly disagree*) and therefore it was combined with category 2 (*Disagree*). The modified category structure can be seen in Table 16.

Table 16. *Category Structure Functioning for Instructional Strategies*

	Count (%)	Infit MNSQ	Outfit MNSQ	Structure calibration	Category measure
1 Disagree	20 (1)	1.20	1.23	None	(-4.16)
2 Slightly disagree	139 (4)	1.05	1.09	-2.91	-2.25
3 Slightly agree	634 (20)	.97	.97	-1.47	-.46
4 Agree	1602 (50)	.94	.97	.41	2.21
5 Strongly agree	789 (25)	.99	.99	3.97	(5.09)

After combining the first two categories the five-point scale met the necessary criteria and performed well based on Linacre’s guidelines, including the .81 logit criterion for separation between adjacent thresholds for a 5-point scale.

The infit and outfit mean square statistics (Table 17) for all items fell between the criterion range of .5-1.50 and no items misfit the model. Part-measure correlations ranged between .60 and .72.

Table 17. *Rasch Item Statistics for Efficacy in Instructional Strategies*

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt-Measure Correlation
IS 3	1.16	.08	1.07	1.0	1.10	1.4	.72
IS 2	.66	.09	.88	-1.6	.93	-.9	.66
IS 7	.49	.09	.92	-1.0	.97	-.4	.73
IS 4	.46	.09	1.32	3.9	1.30	3.7	.69
IS 5	-.01	.09	1.23	2.8	1.23	2.8	.64
IS 6	-.23	.09	.75	-3.4	.76	-3.3	.75
IS 8	-.99	.10	.74	-3.7	.78	-2.7	.69
IS 1	-1.54	.10	.89	-1.5	.92	-.8	.60

A PCA of item residuals was calculated for all IS items, and the items appeared to form a fundamentally unidimensional construct. The amount of variance explained by the Rasch measures was 51.9% (eigenvalue = 8.6), which was slightly more than the suggested value of 50%. The unexplained variance in the first contrast was 8.8% and the eigenvalue was 1.5, which was below the 3.0 criterion. The Rasch item reliability

estimate was .99 and item separation was 8.83 and above the 2.0 benchmark. The Rasch person reliability estimate (.90) and separation (2.02) satisfactorily met their benchmarks.

The Wright map for Instructional Strategies is displayed in Figure 9. The scale ranges from -2 to 7, with an item mean of 0. A comparison of the locations of the means (M) for the person ability measures and item difficulty measures show that the mean person measure ($M = 2.12$; $SD = 1.66$) was higher than the mean item measure ($M = .00$; $SD = .84$). This indicates that the participants easily endorsed the items. This suggests that they believed that they were highly capable of carrying out the instructional strategies described by the items. Item IS 1 (*I can respond effectively to English language questions from my students*; difficulty estimate = -1.54) was the easiest item to endorse, while item IS 3 (*I can provide appropriate support for my least proficient English language students*; difficulty estimate = 1.16) was the most difficult item to endorse. The item hierarchy in Figure 9 suggests that the respondents to the questionnaire felt confident in giving explanations to students and responding to their questions, but less so in providing appropriate support to both low and high level students as well as gauging their students' comprehension and adjusting lessons accordingly. Furthermore, the difficulty hierarchy of items IS 6 (*I can put together appropriate learning activities for my students*) which was easier to endorse than IS 7 (*I can adjust my lessons to different levels for individual students*), which was easier to endorse than IS 2 (*I can gauge my students' comprehension of what I've just taught*), makes sense because each item logically takes more effort than the next (i.e., simply creating or putting an activity together is easier to do than adjusting that activity to perfectly fit individual students,

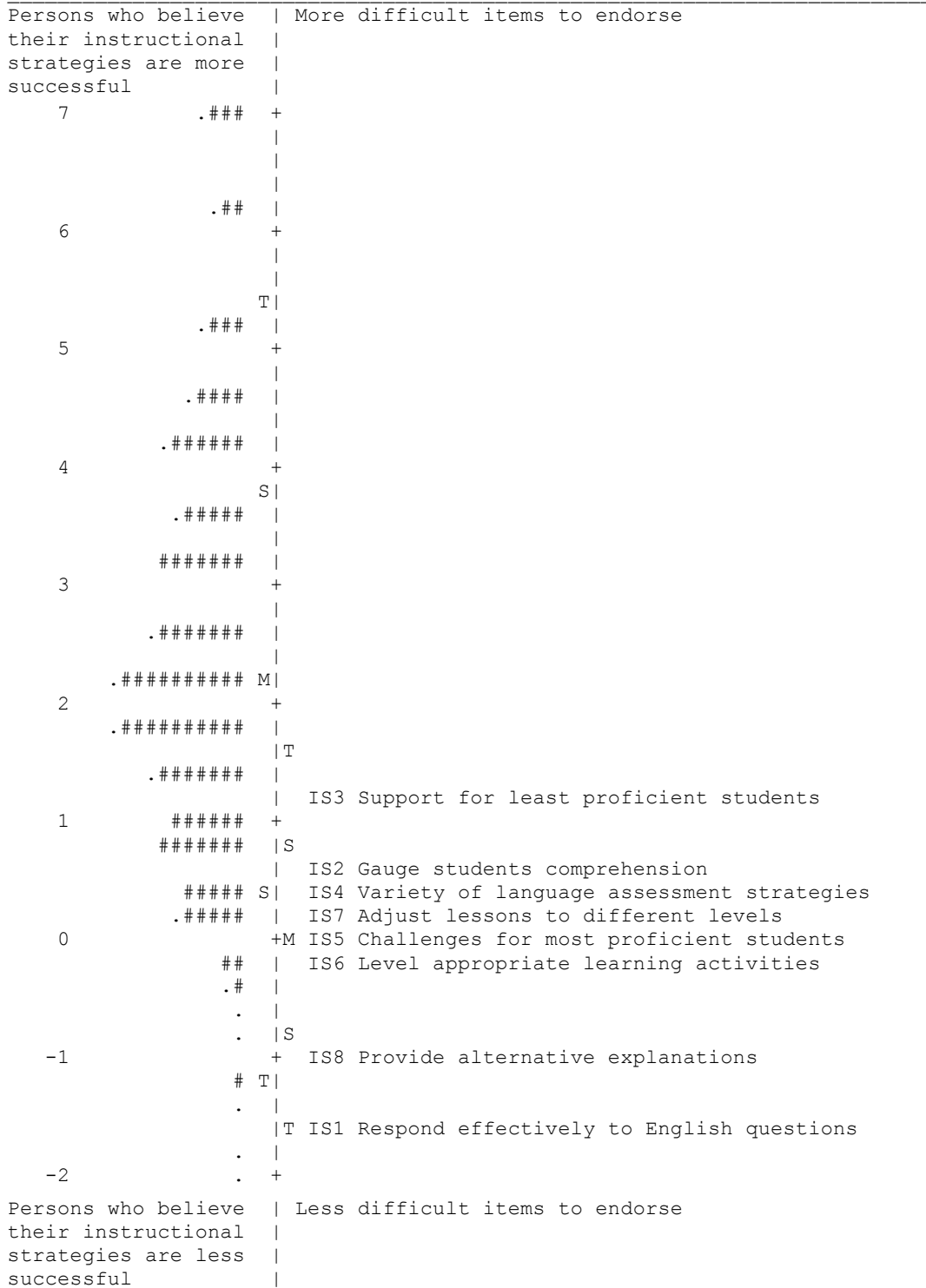


Figure 9. Wright map Instructional Strategies. Each “#” = 4 persons. Each “.” = 1 to 3 persons. M = Mean; S = 1 standard deviation; T = 2 standard deviations.

which is easier to do than gauging if those students actually understood what was taught). Furthermore, the Wright map suggests a ceiling effect in that persons with high ability estimates (i.e., felt confident in their instructional strategies) were not measured as precisely because not enough items were available to measure persons at that level of ability.

In sum, the eight items measuring Efficacy in Instructional Strategies fall between the acceptable criterion range of the infit MNSQ statistic and appear to form a fundamentally unidimensional construct. Furthermore, the variance explained by the Rasch model was above the 50% criterion and the unexplained variance in the first contrast was below the 3.0 eigenvalue benchmark. Item separation and reliability were also both very strong.

Summary

Through the use of the Rasch model, the four hypothesized factors were confirmed to be fundamentally unidimensional. The results are summarized in Table 18.

Table 18. *A Summary of the Four Factors Identified by the Questionnaire*

Name	Label	Item number	Variance accounted for by the Rasch Model (%)	Eigenvalue of the first residual	IR	IS
Dealing with Superiors	DS	2, 3, 5, 6, 7, 8	75.9	1.9	.98	7.68
Student Engagement	SE	1, 2, 3, 4, 5, 6, 7, 8	62.9	1.6	.99	11.17
Classroom Management	CM	1, 2, 3, 4, 5, 6, 7, 8	57.2	1.8	.98	7.45
Instructional Strategies	IS	1, 2, 3, 4, 5, 6, 7, 8	51.9	1.5	.99	8.83

Note. IR = Rasch item reliability; IS = Rasch item separation.

Instrument Validation: Confirmatory Factor Analysis

In addition to the Rasch analysis, a confirmatory factor analysis (CFA) was employed to compare results and further validate the main questionnaire (JULTEBS). As explained in the methodology chapter, the JULTEBS consisted of 32 items hypothesized to measure four self-efficacy constructs: Efficacy in Student Engagement, Efficacy in Instructional Strategies, Efficacy in Classroom Management, and Efficacy in Dealing with Superiors.

The proposed model (Figure 10) shows the configuration of items and constructs included in the JULTEBS. This model represents a typical configuration for a first-order, four-factor structure. The ovals represent the four latent constructs, Student Engagement, Instructional Strategies, Classroom Management, and Dealing with Superiors, while the 32 rectangles represent the 32 questionnaire items, each of which is posited to be composed of a contribution from its respective factor, plus an error term (indicated as the rightmost rectangular 'E' term). For each factor, the metric must be defined by fixing one path, which is indicated by the values 1.0 for items 1, 2, 3, and 4. In addition, all the error paths (not shown) are fixed to 1.0.

Confirmatory Factor Analysis Procedures

The confirmatory factor analysis was conducted using EQS 6.1 software (Bentler, 1995). All data were previously screened, and normality assumptions were confirmed. Model fit was assessed through the use of three commonly reported fit

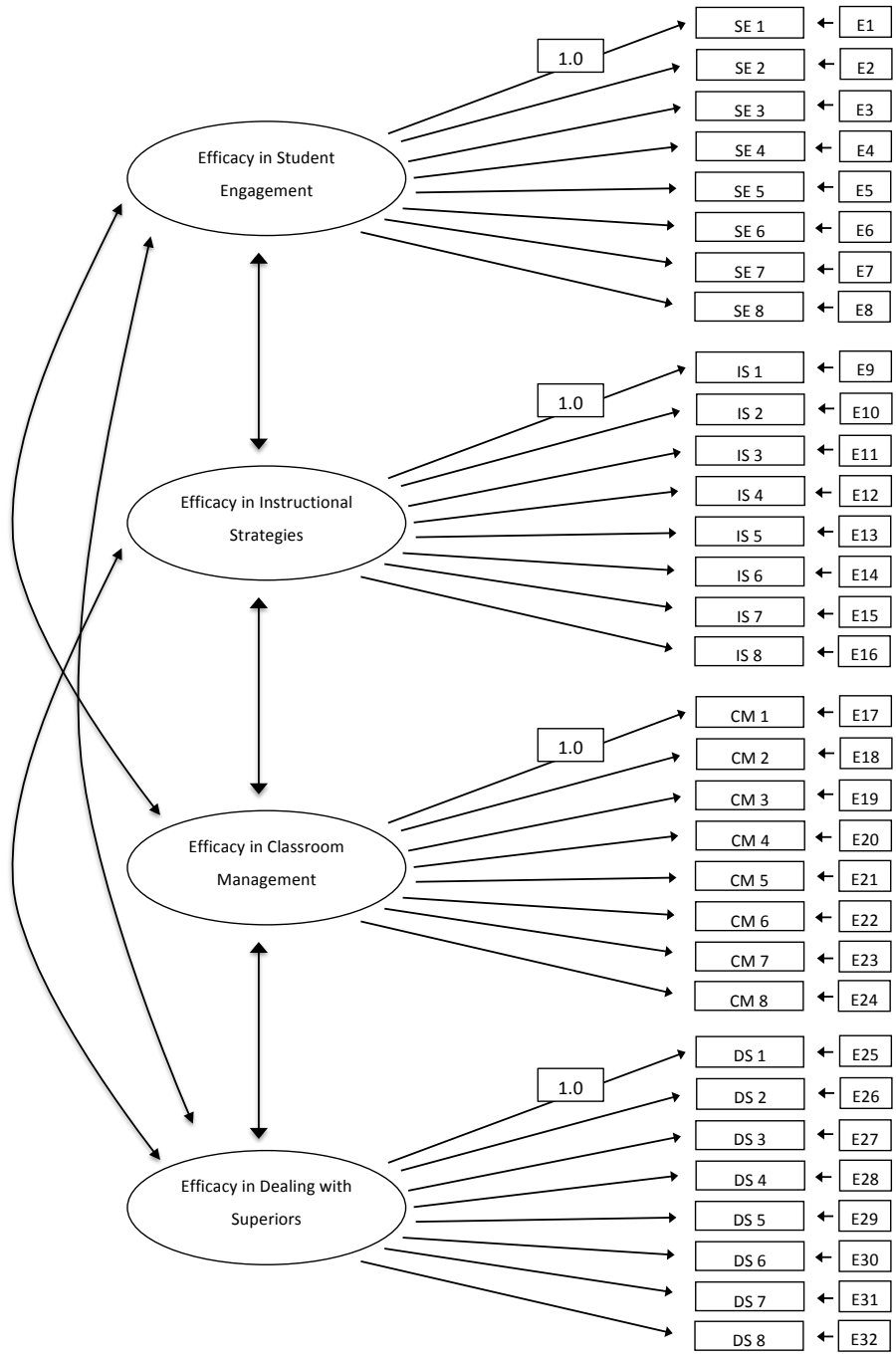


Figure 10. Hypothesized four-factor university English language teacher self-efficacy model.

indices: chi-squared (χ^2), comparative fit index (CFI), and root mean square error of approximation (RMSEA) with 90% confidence intervals.

Confirmatory Factor Analysis Results

As displayed in Figure 11, the proposed model did not display adequate fit: CFI = .885; RMSEA = .07, 90% CIs [.066, .075]; $\chi^2 = 1357.33$, 458 *df*, $p < .01$. Although the RMSEA index met the .08 benchmark for acceptable fit, CFI did not meet the .90 criteria. Based on these results, the model was respecified by deleting several items and cases, as well as correlating errors among items as described in the next section.

Respecification of the Model

In addition to the three fit indices explained above, in order to achieve a better fitting model, the following statistics were also considered in each respecification of the model: the cases (respondents) that contributed the most to multivariate kurtosis, and the standardized coefficients (i.e., the factor loadings). The LaGrange Multiplier (LM) and Wald test results were also consulted to determine if any paths needed to be respecified. Essentially, the LM test begins at the null hypothesis and asks whether movement towards the alternative hypothesis would be an improvement. On the other hand, the Wald test starts at the alternative hypothesis and considers if movement towards the null hypothesis would be an improvement. In other words, the LM test is a general principle for testing hypotheses about parameters in a likelihood framework. To perform an LM test only estimation of the parameters subject to the restrictions is required. In the

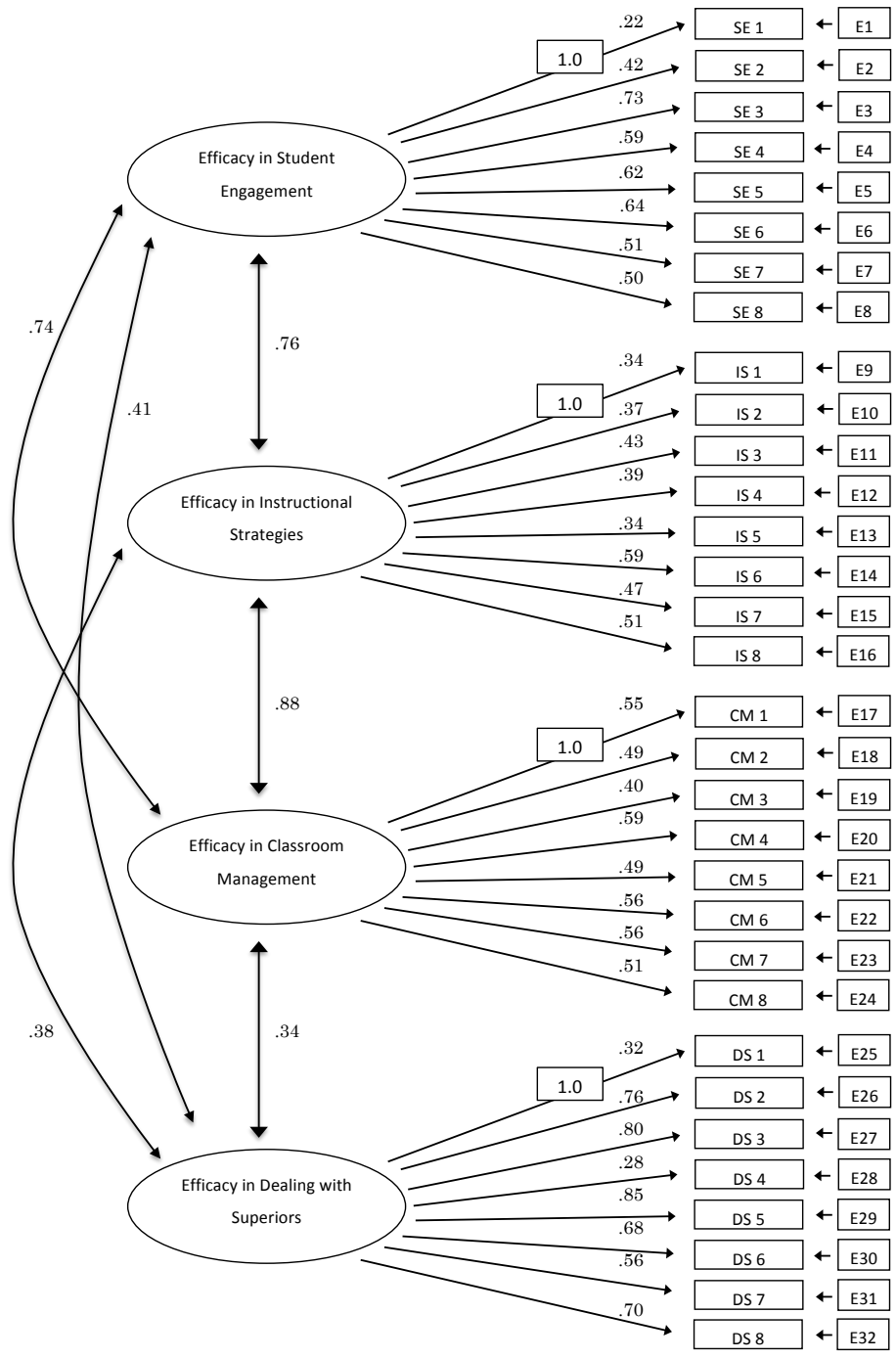


Figure 11. Standardized coefficients for the four-factor self-efficacy model.

context of this study and the four-factor self-efficacy model, the LM test deals with the addition of parameters to the model. This is in contrast with Wald tests, which are based on unrestricted estimates, and likelihood ratio tests, which require both, restricted and unrestricted estimates (i.e., the deletion of parameters from the model). In the current study, only the LM test was used to improve the model as one parameter was added in the final iteration by correlating the errors of e30-e31. The best fitting model was achieved after subjecting the data to five iterations in EQS. In total, five items and six cases were deleted and one parameter was added.

The five items that were deleted were items SE 1, SE 5, SE 7, DS 1, and DS 4. Item SE 1 (*I can communicate in a meaningful way with my students*) showed the weakest loading (.22) of any item on to its hypothesized factor and was therefore deleted. Items DS 1 (*I can communicate my opinions effectively to my superiors*) and DS 4 (*I can avoid being adversely affected by school bureaucracy*) not only had the next two lowest loadings of .32 and .28, respectively, but also misfit in the Rasch analysis, so they were deleted. Finally, items SE 5 (*I can get my students to believe they can succeed in English class*) and SE 7 (*I can foster student creativity*) were deleted based on their high loadings on two different latent variables. These two items loaded not only on to the expected Student Engagement latent variable, but also on to the Instructional Strategies latent variable with standardized coefficients greater than .40. Based on the wording of the two items, SE 5 (*I can get my students to believe they can succeed in English class*) and SE 7 (*I can foster student creativity*), it was conceivable that a logical connection could be made between the items and Instructional Strategies. For example, ‘getting a student to believe’ and ‘fostering student creativity’ could logically be misconstrued as a technique

or a strategy that originates from the side of the teacher. Furthermore, if the items were viewed as teacher centric, it makes sense that they would double load on to Student Engagement and Instructional Strategies latent variables. Therefore, in order to improve model fit and avoid any confusion, the items were deleted instead of adding parameters linking them to two different factors.

In addition to the 42 cases that were eliminated during the data-screening phase, six additional cases (numbers 5, 107, 213, 266, 298, and 313) were deleted due to their excessive contribution to multivariate kurtosis. After the deletions, a total of 392 cases remained. Finally, as mentioned above, based on the results of the LaGrange Multiplier test, one parameter was added to the model by correlating the errors of e30 (*I can get my superiors to support decisions I make regarding English teaching*) to e31 (*I can persuade my superiors to accommodate my teaching needs*), both of which are associated with items measuring Dealing with Superiors.

Lastly, Figure 12 shows the revised and final four-factor self-efficacy model. All of the goodness of fit indices improved and were within the acceptable ranges (CFI = .92; RMSEA = .06, 90% CIs [.055, .066]; $\chi^2 = 775.11, 317df, p < .01$). Aside from the all-round modest to strong loadings displayed between the latent constructs and the items, the loadings between the four latent constructs are also displayed in Figure 12. It is not surprising to see the strong correlations between Student Engagement, Instructional Strategies, and Classroom Management. Not only do all three latent constructs aim to measure the self-efficacy of teacher-student interactions, but also when subjected to a factor analysis during the development of the TSES (Tschannen-Moran & Woolfolk Hoy, 2001), in which this instrument is partly based, similarly high loadings were found. In

addition to the expected strong loadings among Student Engagement, Instructional Strategies, and Classroom Management, it is equally unsurprising that the Dealing with Superiors construct is not as highly correlated with the other three. Dealing with Superiors does not address teacher-student interactions, but is instead designed to measure the self-efficacy of teachers and their relationships with their superiors. Furthermore, this construct was not included in the TSES and was made specifically for the Japanese university English teacher.

Instrument Validation Summary

In this chapter, I presented the instrument validation results derived from the Rasch analysis and confirmatory factor analysis. The four hypothesized factors (Student Engagement, Instructional Strategies, Classroom Management, and Dealing with Superiors) were all found to be unidimensional through the Rasch analysis. In addition, out of the 32 items tested in the Rasch analysis, only two items, DS 1 and DS 4 were found to be problematic.

Despite DS1 and DS4 being problematic in the Rasch analysis, in the CFA model (see Figure 12), all 32 original items were again entered to compare from scratch this analysis to the Rasch analysis. Acceptable fit indices were also achieved and the items loaded onto their respective factors. The CFA did, however, also indicate that five items (SE 1, SE 5, SE 7, DS 1, and DS 4) poorly fit the model.

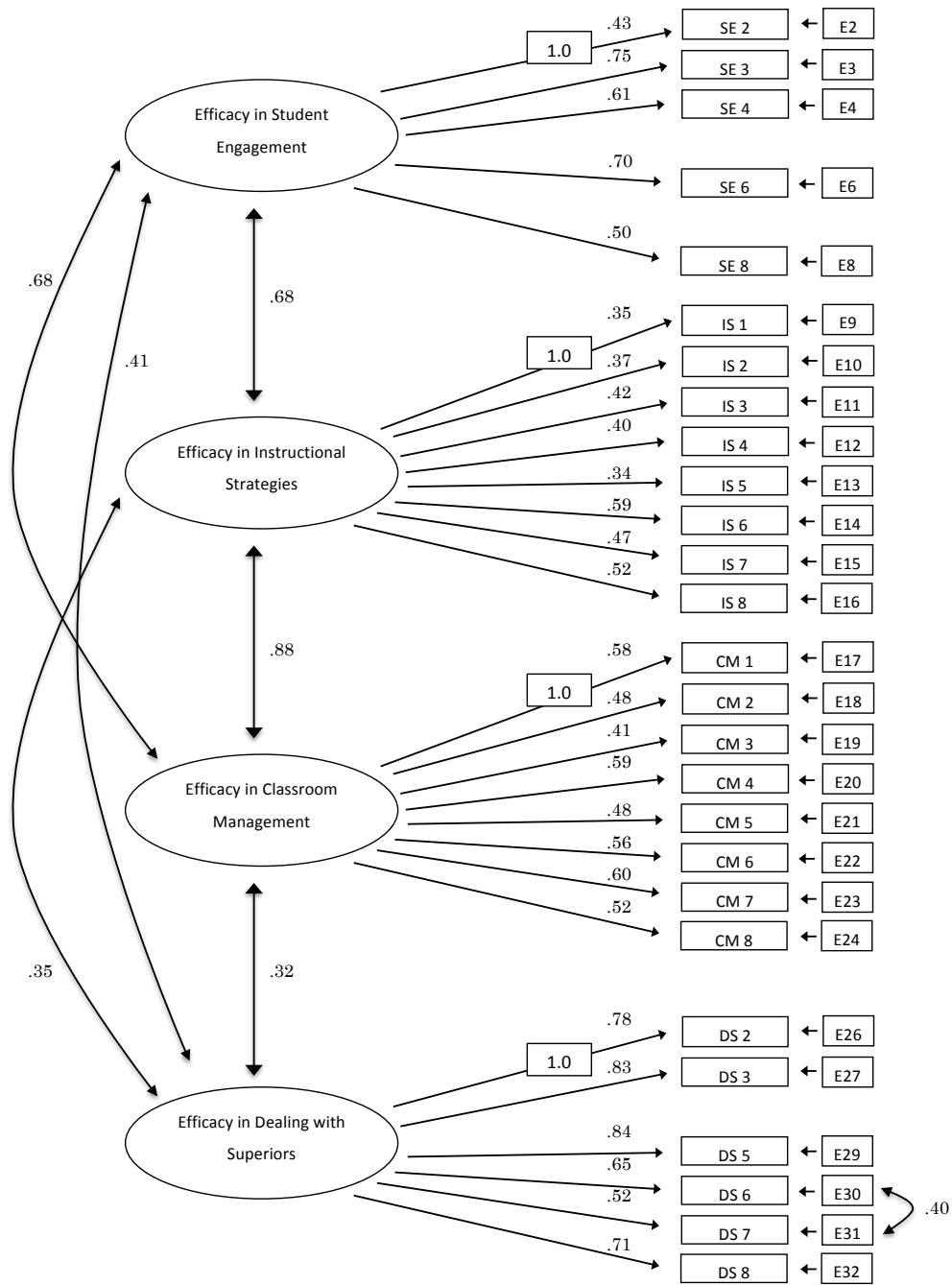


Figure 12. Final version of the four-factor self-efficacy model.

Before moving to the main analysis, a consensus was needed about the five items deemed to be problematic by the two analyses performed above. As was indicated by both the Rasch analysis and the CFA, items DS 1 and DS 4 were problematic and therefore were deleted from the instrument. Items SE 1, SE 5, and SE 7 however were found to only behave poorly in the CFA. The following decisions were made about whether to keep or delete the items in question.

Item SE 1 (*I can communicate in a meaningful way with my students*) was deleted because in the Rasch analysis the infit and outfit MNSQ statistics (1.40, 1.83) were the worst among all the items. Furthermore, this item was extremely easy to endorse. Item SE 7 (*I can foster student creativity*) was also deleted from the instrument. In the confirmatory factor analysis, this item loaded almost equally on both the Student Engagement and Instructional Strategies factors with standardized coefficients of .51 and .50 respectively. Additionally, the wording of this item was deemed to be ambiguous and perhaps too general to accurately assess the respondents' self-efficacy for student engagement. Item SE 5 (*I can get my students to believe they can succeed in English class*) was kept because it not only had a .62 loading on to the SE factor, but out of the three items in question, it logically made the most sense because it directly inquires about engaging one's students and thus fit the Student Engagement construct best.

To conclude this chapter, Tables 19 through 21 provide a visual summary of the post-modification remaining data to be used in the statistical procedures in Chapter 5.

Table 19. *A Summary of the Final Number of Respondents*

Original number of respondents	Number of respondents after the initial data screening	Final number of respondents after the deletion of cases contributing to multivariate kurtosis in the CFA
440	398	392

Table 20. *A Summary of the Final Questionnaire Items to be used in the Main Analysis*

Name	Label	Item number
Student Engagement	SE	2, 3, 4, 5, 6, 8
Instructional Strategies	IS	1, 2, 3, 4, 5, 6, 7, 8
Classroom Management	CM	1, 2, 3, 4, 5, 6, 7, 8
Dealing with Superiors	DS	2, 3, 5, 6, 7, 8

Table 21. The Final *Japanese University Language Teacher Efficacy Beliefs Scale (JULTEBS)*

Factors	JULTEBS items
Student Engagement	SE 2. I can help my students think critically in English.
	SE 3. I can motivate most of my students to become more interested in English.
	SE 4. I can help students become more independent English learners.
	SE 5. I can get my students to believe they can succeed in English class.
	SE 6. I can get most of my students to value learning English.
Instructional Strategies	SE 8. I can greatly improve the overall English skills of my students.
	IS 1. I can respond effectively to English language questions from my students.
	IS 2. I can gauge my students' comprehension of what I have just taught.
	IS 3. I can provide appropriate support for my least proficient English language students.
	IS 4. I can use a variety of language assessment strategies in my classes.
	IS 5. I can provide appropriate challenges for my most proficient English language students.
	IS 6. I can put together level appropriate learning activities for my students.
	IS 7. I can adjust my lessons to different levels for individual students.
IS 8. I can provide alternative explanations when students are confused.	
Classroom Management	CM 1. I can control disruptive behavior in my English classroom.
	CM 2. I can make my expectations clear about appropriate student behavior.
	CM 3. I can create a learning friendly environment in my classroom.
	CM 4. I can deal effectively with defiant students.
	CM 5. I can establish effective routines in my classes.
	CM 6. I can get my students to follow classroom rules.
	CM 7. I can establish my own classroom management system.
	CM 8. I can prevent problem students from ruining classes.
Dealing with Superiors	DS 2. I can convince my superiors to listen to my suggestions.
	DS 3. I can convince my superiors to act on my suggestions.
	DS 5. I can convince my superiors to take my suggestions seriously.
	DS 6. I can get my superiors to support decisions I make regarding English teaching.
	DS 7. I can persuade my superiors to accommodate my teaching needs.
	DS 8. I can convince my superiors to have greater vision regarding language teaching.

Note. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors.

CHAPTER 5

QUANTITATIVE RESULTS

The purpose of this chapter is to present the quantitative results for the first four research questions. After describing the initial data screening, Rasch analysis, and confirmatory factor analysis in the previous chapter, four profile analyses are presented in this chapter using the Rasch logit person ability measures that were transformed from the scores obtained on the self-efficacy questionnaire.

Research Question 1: Self-Efficacy and Native Language

The first research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' native language. This question is answered by conducting a profile analysis, where the teacher's native language is entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors are entered as the dependent variables.

The profile analysis checks if the groups (a) exhibit flatness across observations, (b) have parallel profiles between observations, and (c) have equal levels across observations. In other words, the flatness of the profiles checks the similarity of responses to all dependent variables, independent of groups. The parallelism, an interaction term, test checks whether different groups have similar profiles and whether interaction effects exist among different groups. Lastly, the levels hypothesis checks whether one group, on average, scores higher on the collected set of measures than another group.

Table 22 provides the descriptive statistics of the data used in the first profile analysis.

Table 22. *Descriptive Statistics for Native and Non-Native English Teachers Across the Four Self-Efficacy Factors*

Group	Factor	<i>M</i>	<i>SE</i>	95% CI	<i>SD</i>
Native Teachers (<i>n</i> = 209)	SE	2.67	.12	[2.43, 2.91]	1.76
	IS	3.18	.11	[2.97, 3.34]	1.52
	CM	4.78	.18	[4.43, 5.12]	2.53
	DS	1.41	.15	[1.11, 1.70]	2.15
Japanese Teachers (<i>n</i> = 183)	SE	1.98	.13	[1.72, 2.23]	1.76
	IS	2.57	.11	[2.36, 2.78]	1.46
	CM	3.47	.17	[3.14, 3.80]	2.26
	DS	1.21	.15	[0.92, 1.50]	2.01

Note. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors. All statistics are based on Rasch person ability estimates.

Table 23 shows the flatness of the four self-efficacy factors, the parallelism for the four self-efficacy factors and the two groups of native English and Japanese teachers, and the differences of the levels between the two groups. All three contrasts in the profile analysis were significant.

Table 23. *Profile Analysis for the Four Self-Efficacy Factors by Native Speaker Group*

Source of Variance	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η^2	<i>Power</i>
Within Group							
Self-efficacy (flatness)	1607.26	3	696.76	172.75	.00	.57	1.00
Self-efficacy * Group (parallelism)	61.57	3	26.69	7.18	.00	.05	.98
Error	2307.99	1170	2.57				
Between Group							
Group (level)	191.78	1	191.78	20.05	.00	.05	.99
Error	3730.28	390	9.57				

Note. GLM offers four types of tests for the parallelism, levels, and flatness tests: Sphericity Assumed, Greenhouse-Geisser, Huynh-Feld, and lower-bound. As the Sphericity assumption was not met, the Greenhouse-Geisser correction (the most conservative) statistics are given.

The first effect shown in Table 23 is the self-efficacy effect for flatness and it indicates that the results deviated significantly from flatness, $F = 172.75$, $p < .001$, partial $\eta^2 = .57$, power = 1.00. As you can see in Figure 12, teachers' self-efficacy varied significantly regardless of whether they were a native English teacher or a native Japanese English teacher. In other words, irrespective of whether the teacher was a native English speaker or a native Japanese speaker, their self-efficacy levels went up and down and were different for each self-efficacy factor (i.e., the profiles were not flat for either group).

The second effect shown in Table 23 is the interaction effect for parallelism, and it indicates that the native English teachers and Japanese English teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared, $F = 7.18$, $p < .001$, partial $\eta^2 = .05$, power = .98.

The third effect shown in Table 23 is the group effect for level. It indicates significant differences between groups, where teachers' self-efficacy varied significantly according to their native English speaker or native Japanese speaker status, $F = 20.05$, $p < .001$, partial $\eta^2 = .05$, power = .99. Native English teachers' self-ratings were significantly higher than those for the native Japanese teachers. In other words, when the groups as a whole were compared, the native English teachers profiles were higher than their Japanese counterparts.

Finally, Figure 12 displays the profiles of the native English teachers and the Japanese English teachers across the four self-efficacy factors (Student Engagement, Instructional Strategies, Classroom Management, and Dealing with Superiors). The four

factors are on the x-axis, and the Rasch person ability estimates from the self-efficacy questionnaire are on the y-axis. It is evident that the Native English teachers have higher Rasch person ability estimates than those of the Japanese for almost every self-efficacy factor.

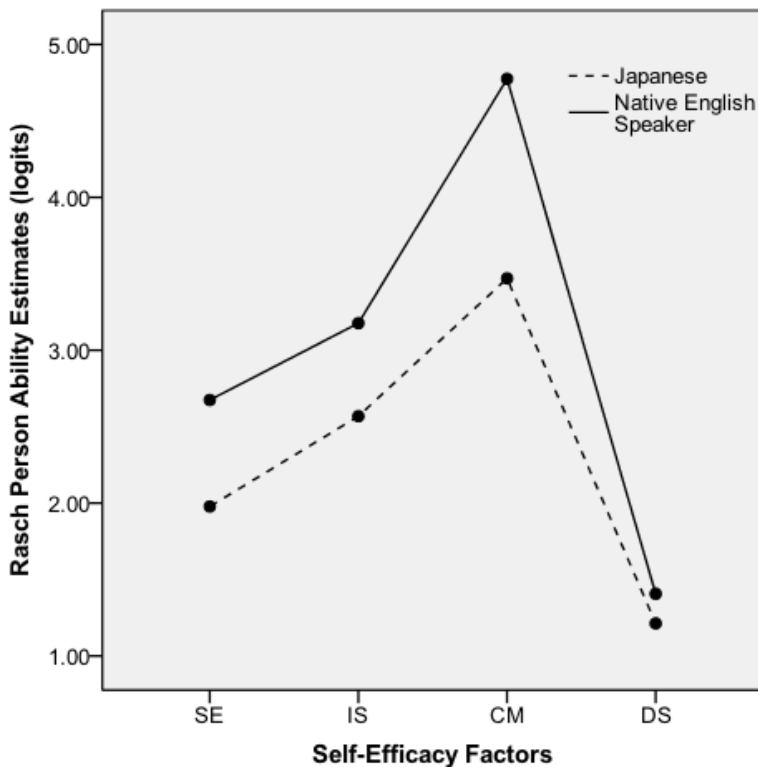


Figure 13. Profiles of the four self-efficacy factors for the native English teachers and Japanese English teachers. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors.

Research Question 2: Self-Efficacy and Teaching Experience

The second research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' length of teaching experience. The teachers are split into three groups: 0 to 10 years of experience, 11 to 20

years of experience, and more than 20 years of experience. This question is answered by conducting a profile analysis, where the teacher's experience is entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors are entered as the dependent variables.

Table 24 provides the descriptive statistics of the data used in the second profile analysis.

Table 24. *Descriptive Statistics for Teaching Experience Across the Four Self-efficacy Factors*

Group	Factor	<i>M</i>	<i>SE</i>	95% CI	<i>SD</i>
0-10 Years (<i>n</i> = 114)	SE	2.13	.13	[1.87, 2.40]	1.44
	IS	2.41	.12	[2.17, 2.66]	1.31
	CM	3.61	.21	[3.20, 4.01]	2.20
	DS	1.10	.19	[0.67, 1.43]	2.10
11-20 Years (<i>n</i> = 166)	SE	2.34	.15	[2.05, 2.64]	1.91
	IS	3.03	.13	[2.78, 3.28]	1.62
	CM	4.36	.20	[3.96, 4.75]	2.60
	DS	1.22	.18	[0.88, 1.57]	2.26
20+ Years (<i>n</i> = 112)	SE	2.58	.18	[2.22, 2.94]	1.90
	IS	3.18	.14	[2.90, 3.45]	1.47
	CM	4.45	.24	[3.98, 4.93]	2.55
	DS	1.72	.17	[1.39, 2.05]	1.77

Note. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors. All statistics are based on Rasch person ability estimates.

Table 25 shows the flatness of the four self-efficacy factors, the parallelism for the four self-efficacy factors and the three groups of teachers, and the differences of the levels between the three groups. Two of the three contrasts (the flatness test and levels test) in the profile analysis were significant.

Table 25. Profile Analysis for the Four Self-Efficacy Factors by Experience

Source of Variance	SS	df	MS	F	p	η^2	Power
Within Group							
Self-efficacy (flatness)	1550.61	3	680.98	161.42	.00	.56	1.00
Self-efficacy * Group (parallelism)	19.66	6	4.32	1.62	.14	.01	.62
Error	2349.90	1167	2.01				
Between Group							
Group (level)	109.43	2	54.71	5.58	.00	.03	.86
Error	3812.63	389	9.80				

Note. GLM offers four types of tests for the parallelism, levels, and flatness tests by default: Sphericity Assumed, Greenhouse-Geisser, Huynh-Feld, and Lower-bound. As the Sphericity assumption was not met, the Greenhouse-Geisser correction (the most conservative) statistics are given.

The first effect shown in Table 25 is the self-efficacy effect for flatness and it indicates that the results deviated significantly from flatness, $F = 161.42$, $p < .001$, partial $\eta^2 = .56$, power = 1.00. In other words, teachers' self-efficacy varied significantly regardless of the number of years teaching English.

The second effect shown in Table 25 is the interaction effect for parallelism, and it indicates that the three levels of experienced teachers did not significantly show different patterns of highs and lows when the means of the self-efficacy scores were compared, $F = 1.62$, $p > .001$, partial $\eta^2 = .01$, power = .62.

The third effect shown in Table 25 is the group effect for level. It indicates significant differences between the three groups, where teachers' self-efficacy varied significantly according to the number of years teaching university English, $F = 5.58$, $p < .001$, partial $\eta^2 = .03$, power = .86.

Because both the flatness and levels test showed statistically significant differences, a simple effects analysis was conducted in which the differences among

group means were examined separately for each factor. In order to pinpoint where the differences were, a univariate follow-up one-way ANOVAs with a Scheffé post hoc test were performed using the three experience groups as the independent variable and the four factors as dependent variables. The results indicated that a significant mean difference exists between the 0-10 year experience group and the 11-20 year group ($p < .05$) as well as between the 0-10 year group and the 20+ years group ($p < .05$) for both the Instructional Strategies and the Classroom Management variables.

Finally, Figure 13 displays the profiles of the three groups of teachers with different amounts of experience across the four self-efficacy factors. The four factors are

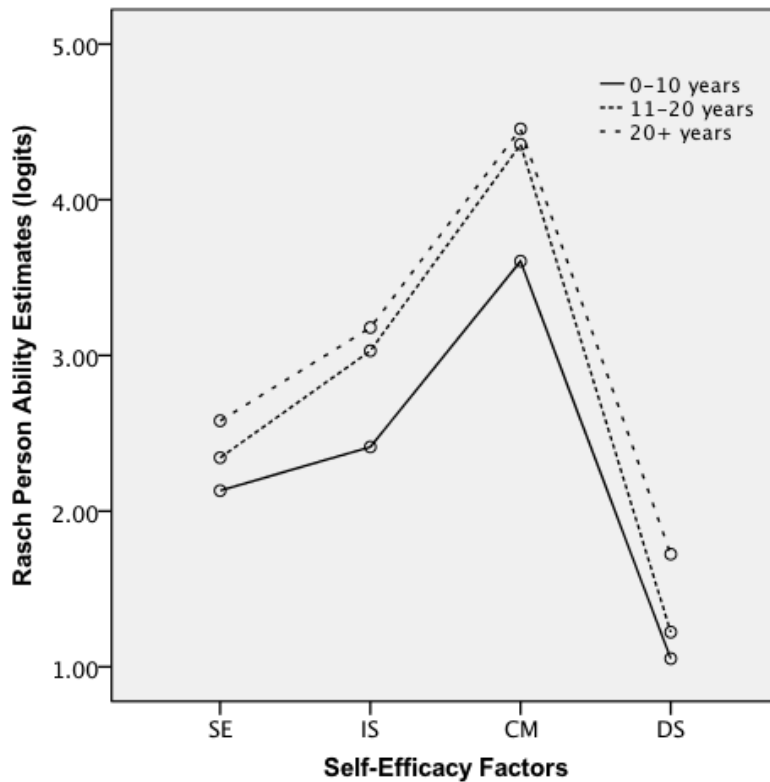


Figure 14. Profiles of the four self-efficacy factors for the teachers' years of experience. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors.

on the x-axis, and the Rasch person ability estimates from the self-efficacy questionnaire are on the y-axis. It is clear that while the more experienced teachers (i.e., 11-20 years and 20+ years) had similar Rasch person ability estimates, the less experienced teachers (i.e., 0-10 years) showed much lower ability estimates for every self-efficacy factor.

Research Question 3: Self-Efficacy and Contract Status

The third research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' tenured or term-limited contract status. This question is answered by conducting a profile analysis, where the teacher's employment status is entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors are entered as the dependent variables.

Table 26 provides the descriptive statistics of the data used in the third profile analysis.

Table 26. *Descriptive Statistics for Tenured and Limited-Term Contract English Teachers Across the Four Self-Efficacy Factors*

Group	Factor	<i>M</i>	<i>SE</i>	95% CI	<i>SD</i>
Tenured Teachers (<i>n</i> = 172)	SE	2.25	.14	[1.98, 2.52]	1.80
	IS	2.84	.11	[2.61, 3.06]	1.49
	CM	3.94	.18	[3.58, 4.29]	2.38
	DS	1.73	.15	[1.44, 2.03]	1.97
Contract Teachers (<i>n</i> = 220)	SE	2.43	.12	[2.19, 2.66]	1.78
	IS	2.94	.10	[2.73, 3.14]	1.55
	CM	4.35	.17	[4.00, 4.67]	2.57
	DS	.99	.14	[0.70, 1.27]	2.12

Note. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors. All statistics are based on Rasch person ability estimates.

Table 27 shows the flatness of the four self-efficacy factors, the parallelism for the four self-efficacy factors and the two groups of teachers, and the differences of the levels between the two groups. Two of the three contrasts (the flatness test and parallelism test) in the profile analysis were significant. The first effect shown in Table 27 is the self-efficacy effect for flatness and it indicates that the results deviated significantly from flatness, $F = 166.77$, $p < .001$, partial $\eta^2 = .56$, power = 1.00. In other words, teachers' self-efficacy varied significantly regardless of whether they were a tenured or a term-limited contract teacher.

Table 27. Profile Analysis for the Four Self-Efficacy Factors by Employment Status

Source of Variance	SS	df	MS	F	p	η^2	Power
Within Group							
Self-efficacy (flatness)	1555.77	3	669.84	166.77	.00	.56	1.00
Self-efficacy * Group (parallelism)	74.65	3	32.14	7.66	.00	.06	.99
Error	2294.10	1170	1.96				
Between Group							
Group (level)	.11	1	.11	.01	.92	.00	.05
Error	3921.96	390	10.06				

Note. GLM offers four types of tests for the parallelism, levels, and flatness tests by default: Sphericity Assumed, Greenhouse-Geisser, Huynh-Feld, and Lower-bound. As the Sphericity assumption was not met, the Greenhouse-Geisser correction (the most conservative) statistics are given.

The second effect shown in Table 27 is the interaction effect for parallelism. It indicates that the tenured teachers and the term-limited contract teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared, $F = 7.66$, $p < .001$, partial $\eta^2 = .06$, power = .99.

The third effect shown in Table 27 is the group effect for level. It indicates a lack of significant differences between groups, where teachers' self-efficacy did not vary

significantly according to the teachers' tenured or contract status, $F = .01, p > .001$, partial $\eta^2 = .00$, power = .05.

Finally, Figure 14 displays the profiles of the four self-efficacy factors for the tenured teachers and the term-limited contract teachers. The four self-efficacy factors are on the x-axis, and the Rasch person ability estimates from the self-efficacy questionnaire are on the y-axis. It is clear that while the tenured and contract teachers had similar Rasch person ability estimates and trajectories over the first three factors, the contract teachers showed much lower Rasch person ability estimates for the Dealing with Superiors factor.

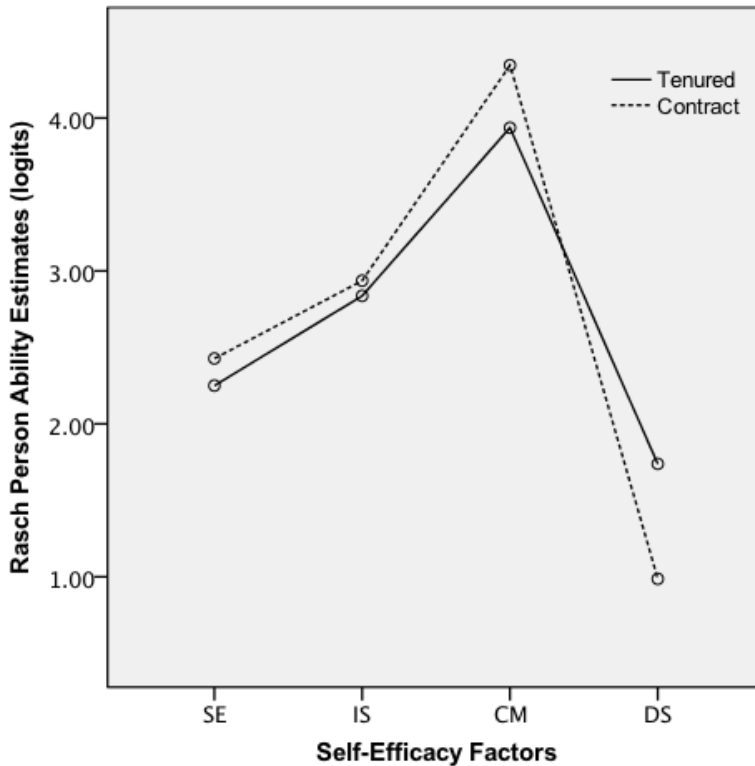


Figure 15. Profiles of the four self-efficacy factors for the teachers' employment status. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors.

Research Question 4: Self-Efficacy and Gender

The fourth research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' gender. This question is answered by conducting a profile analysis, where the teacher's gender is entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors are entered as the dependent variables.

Table 28 provides the descriptive statistics of the data used in the fourth profile analysis.

Table 28. *Descriptive Statistics for Male and Female English Teachers Across the Four Self-Efficacy Factors*

Group	Factor	<i>M</i>	<i>SE</i>	95% CI	<i>SD</i>
Male Teachers (<i>n</i> = 226)	SE	2.43	.12	[2.19, 2.67]	1.81
	IS	2.87	.10	[2.68, 3.07]	1.50
	CM	4.29	.16	[3.96, 4.61]	2.48
	DS	1.31	.13	[1.05, 1.58]	2.02
Female Teachers (<i>n</i> = 166)	SE	2.24	.14	[1.97, 2.51]	1.75
	IS	2.92	.12	[2.68, 3.16]	1.56
	CM	4.00	.20	[3.61, 4.39]	2.51
	DS	1.32	.17	[0.99, 1.65]	2.18

Note. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors. All statistics are based on Rasch person ability estimates.

Table 29 shows the flatness of the four self-efficacy factors, the parallelism for the four self-efficacy factors and the two gender groups, and the differences of the levels between the two groups. Only one of the three contrasts (flatness test) in the profile analysis was significant.

Table 29. Profile Analysis for the Four Self-Efficacy Factors by Gender

Source of Variance	SS	df	MS	F	p	η^2	Power
Within Group							
Self-efficacy (flatness)	1595.19	3	701.71	165.02	.00	.56	1.00
Self-efficacy * Group (parallelism)	7.31	3	3.22	1.76	.15	.01	.46
Error	2362.24	1170	2.02				
Between Group							
Group (level)	4.24	1	4.24	.42	.52	.00	.09
Error	3917.83	390	10.05				

Note. GLM offers four types of tests for the parallelism, levels, and flatness tests by default: Sphericity Assumed, Greenhouse-Geisser, Huynh-Feld, and Lower-bound. As the Sphericity assumption was not met, the Greenhouse-Geisser correction (the most conservative) statistics are given.

The first effect shown in Table 29 is the self-efficacy effect for flatness and it indicates that the results deviated significantly from flatness, $F = 165.02$, $p < .001$, partial $\eta^2 = .56$, power = 1.00. In other words, teachers' self-efficacy varied significantly regardless of whether they were a male or female teacher.

The second effect shown in Table 29 is the interaction effect for parallelism. It indicates that the male teachers and the female teachers did not have significantly different patterns of highs and lows when the means of the self-efficacy scores were compared, $F = 1.76$, $p > .001$, partial $\eta^2 = .01$, power = .46.

The third effect shown in Table 29 is the group effect for level. It indicates a lack of significant differences between groups, where teachers' self-efficacy did not vary significantly according to the teachers' gender, $F = .42$, $p > .001$, partial $\eta^2 = .00$, power = .09.

Finally, Figure 15 displays the profiles of the four self-efficacy factors for the male teachers and the female teachers. The four self-efficacy factors are on the x-axis,

and the Rasch person ability estimates from the self-efficacy questionnaire are on the y-axis. It is evident from the Figure below that both the male and female teachers had extremely similar Rasch person ability estimates and trajectories over the all four self-efficacy factors.

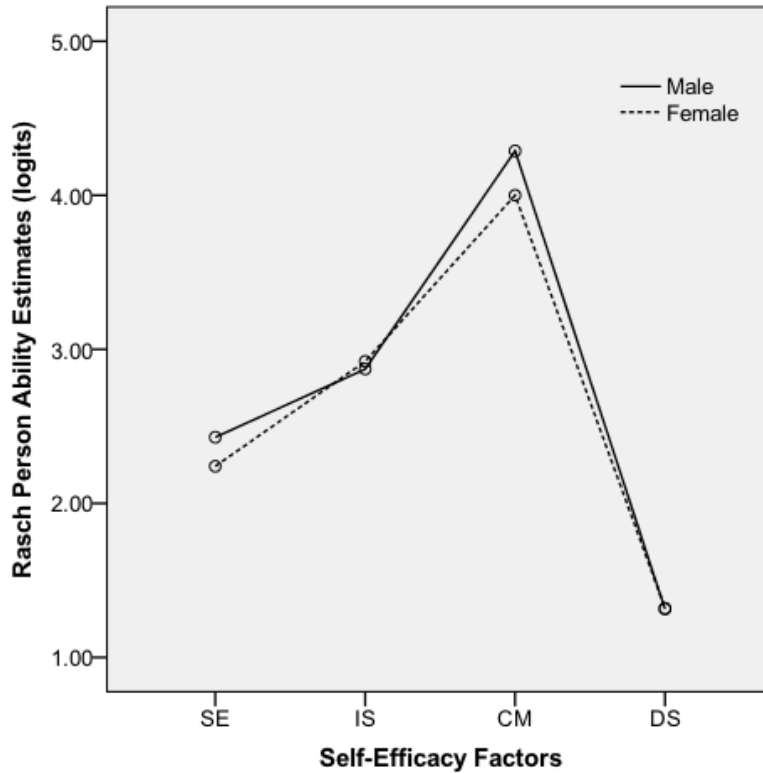


Figure 16. Profiles of the four self-efficacy factors for the teachers' gender. SE = Student Engagement; IS = Instructional Strategies; CM = Classroom Management; DS = Dealing with Superiors.

Summary

In this chapter, the quantitative results of four profile analyses were reported to answer the first four research questions. Table 30 shows a summary of the results from the three primary tests performed in a profile analysis for each independent variable.

Furthermore, the patterns shown in Figures 12 to 15 are also noteworthy. Specifically, all

of the Figures in this chapter show an identical trend, where the Classroom Management factor is always the highest and the Dealing with Superiors factor is always the lowest. Also, the Instructional Strategies factor was always higher than the Student Engagement factor. These trends and patterns will be considered in greater detail in Chapter 7.

Table 30. *Summary of the Profile Analyses Results*

Independent variable	Test of flatness	Test of parallelism	Test of levels
Native language	Significant	Significant	Significant
Teaching experience	Significant	Non-significant	Significant
Contract status	Significant	Significant	Non-significant
Gender	Significant	Non-significant	Non-significant

Note. Significance is measured at $p < .001$.

In the next chapter, I report the qualitative results in order to illuminate how the development of self-efficacy beliefs of 12 university English teachers in Japan is supported and challenged.

CHAPTER 6

QUALITATIVE RESULTS AND DISCUSSION

The purpose of this chapter is to report the results of the qualitative data, which were obtained through semi-structured interviews conducted with 12 English language teachers currently teaching at Japanese universities. The teachers were selected using a stratified purposeful sampling method. Stratified purposeful sampling is a non-random sampling strategy, where the researcher selects information-rich cases for in-depth study (Patton, 2001). Information-rich cases are those from which one can learn a great deal about issues of central importance to the purpose of the research. In the current study, such cases were teachers with various backgrounds, ages, and experiences who were willing to speak with me candidly about teaching English at a Japanese university. This sampling strategy allowed me to help illustrate characteristics of particular subgroups of interest and facilitate comparison between different groups. Finally, as the present study adopted a triangulation strategy mixed methods design (Creswell, 2009), the results of the qualitative data discussed in this chapter were analyzed separately from the quantitative data for the purposes of answering the fifth (What potentially supports strong language teacher self-efficacy beliefs?) and sixth (What potentially challenges language teacher self-efficacy beliefs?) research questions.

Interviews

Participant Characteristics

There were eight male and four female university teachers who were interviewed.

Seven teachers were tenured (i.e., they hold permanent positions until retirement age, usually 60 or 65 years old depending on the university) and five were on non-renewable limited-term contracts (see participant information below for the exact length of contract). Seven were native Japanese speakers and five were native English speakers. Finally, two teachers had between zero and ten years of English teaching experience in Japan, seven teachers had ten to twenty years of experience, and three teachers had more than twenty years of experience. Table 31 displays the bio-information of the teachers who participated in the interviews. Under Table 31 is a more detailed description of each participant and their specific teaching situation. The names listed are pseudonyms to protect the identities of the participants.

Table 31. *Teacher Bio-Information*

Name	Gender	Tenured	Nationality	Experience
Robert	Male	No	American	10-20 years
William	Male	Yes	American	10-20 years
Matt	Male	No	Australian	10-20 years
Kenta	Male	Yes	Japanese	20 + years
Masato	Male	No	Japanese	20 + years
Akira	Male	Yes	Japanese	10-20 years
Kohei	Male	Yes	Japanese	10-20 years
Yoshi	Male	Yes	Japanese	0-10 years
Kate	Female	Yes	American	20 + years
Michelle	Female	No	American	10-20 years
Yuka	Female	Yes	Japanese	10-20 years
Rie	Female	No	Japanese	0-10 years

Robert is a 45 year-old American Associate Professor working at a prestigious National University in eastern Japan. He is in year four of a 6-year limited-term contract and has a total of 14 years of teaching experience. He is stationed at the Language Center within his

university and teaches students from several different departments with various levels of English and motivation. He mainly teaches first and second year students their required English courses.

William is a 42 year-old American tenured Associate Professor working at a large mid-level private university in western Japan. He has been at his current job for 2 years and he has a total of 11 years of teaching experience. He teaches in the Business department to low and mid-level students. He teaches first and second year students their required English courses as well as elective courses to third and fourth year students.

Matt is a 37 year-old Australian lecturer working at a mid-level private university in eastern Japan. He is in year two of a 6-year limited-term contract and has a total of 12 years of teaching experience. He is stationed at the Language Center at his university and teaches students from several different departments with various levels of English and motivation. He mainly teaches first and second year students their required English courses.

Kenta is a 57 year-old Japanese tenured Professor working at a well-respected private university in western Japan. He has been at his current job for 20 years and he has a total of 25 years of teaching experience. He teaches in the Information Science and Technology department to mid-level students. He occasionally teaches first and second year students their required English courses, but mostly upper level elective courses to third and fourth year students.

Masato is a 65 year-old Japanese lecturer working at a well known high-level private university in eastern Japan. He is in year three of a 5-year limited-term contract and has a total of 27 years of teaching experience. Before becoming a teacher, Masato worked as an engineer in Tokyo. He is stationed in the department of Science and Engineering and teaches students with various levels of English and motivation. He mainly teaches first and second year students their required English courses.

Akira is a 38 year-old Japanese tenured Associate Professor working at a very well respected private university in western Japan. He has been at his current job for 3 years and has a total of 12 years of teaching experience. He teaches in the Global Communications department to mid and high level students. He teaches both required English courses as well as elective courses.

Kohei is a 45 year-old Japanese tenured Professor working at a mid-level private university in eastern Japan. He has been at his current job for 5 years and he has a total of 17 years of teaching experience. He teaches in the Economics department to mid-level students. He teaches both first and second year students their required English courses as well as upper level elective courses to third and fourth year students.

Yoshi is a 33 year-old Japanese tenured Associate Professor working at a low level private university in eastern Japan. He has been at his current job for 6 years and has a

total of 6 years of teaching experience. He teaches in the Global Media Studies department to low proficiency students. He teaches both required English courses as well as elective courses.

Kate is a 48 year-old American tenured Professor working at a small mid-level private university in western Japan. She has been at her current job for 18 years and has a total of 23 years of teaching experience. She teaches in the International English department to low, mid, and high level students. She teaches first and second year students their required English courses as well as seminars to third and fourth year students.

Michelle is a 40 year-old American lecturer working at a mid-level private university in western Japan. She is in year two of a four-year limited-term contract and has a total of 10 years of teaching experience. She is stationed at the Language Center and teaches students from several different departments with various levels of English and motivation. She mainly teaches first and second year students their required English courses.

Yuka is a 43 year-old Japanese tenured Associate Professor working at a prestigious National University in eastern Japan. She has been at her current job for 11 years and has a total of 16 years of teaching experience. She teaches in the Humanities department to low, mid, and high-level students. She teaches first and second year students their required English courses as well as seminars to third and fourth year students.

Rie is a 32 year-old Japanese lecturer working at a mid-level private university in western Japan. She is in year two of a four-year limited-term contract and has a total of five years of teaching experience. She is stationed at the Language Center and teaches students from several different departments with various levels of English and motivation. She mainly teaches first and second year students their required English courses.

Patterns and Themes

The interviews were conducted in English, digitally recorded, and lasted between 45 and 90 minutes (see Chapter 3 for a full description). The digital recordings were transcribed and patterns and themes found in the data providing insight in to research questions 5 and 6 were identified and analyzed using coding methods outlined by Tesch (1990). Following this coding process, the identified themes and categories were investigated. Specifically, four main categories were found to address research question 5, regarding what potentially supports the formation of strong language teacher self-

efficacy beliefs. There were three main categories found to help answer question 6, regarding what potentially challenges language teachers' self-efficacy beliefs. Table 32 presents the seven categories found in the interview data. Finally, the results of the data analysis are presented in this chapter by individually discussing each category for both research questions.

Table 32. *Main Categories from the Interview Data*

	Category
RQ 5 Supports teacher self-efficacy	Autonomy Colleagues Money Students
RQ 6 Challenges teacher self-efficacy	Administration Students Contract Teachers

Research Question 5: Potential Self-Efficacy Supports

Autonomy

Several teachers pointed out that one of the most positive aspects of their jobs was the autonomy afforded to them by their employers. They mentioned that being given freedom of choice, particularly concerning their teaching styles in the English classroom, curriculum design, as well as in general decision making processes, gave them a sense of empowerment. Moreover, being given autonomy and then being held responsible as well as accountable for their work was viewed as a positive motivator for most of the teachers, not a burden.

Teaching style.

In relation to teaching styles and autonomy in the classroom, Yuka, a tenured associate professor at a national university said:

I think the last couple of years my teaching has really gotten better. I know my own teaching style and I know how to conduct my own classes. I'm pretty good at keeping the students' attention from the beginning to the end. And the best part is that the school lets me run the classes my way. I can bring my own personality to the classroom, and that's ok with them.

Similarly, Michelle, an American contract teacher working at a private university commented:

One of the best parts of my job is preparing the lessons because I enjoy it so much. I like the creative aspect of planning lessons from scratch and seeing if they succeed or totally flop. If they succeed, that's great. If they don't, it's back to the drawing board. But either way, they don't tell me how to teach. The university allows me to run my own classes in the manner that I choose. This lets me be imaginative, which is a good thing.

One of the commonalities in the quotations above is how both Yuka and Michelle place importance in running their classes in a style that they have chosen. They both are confident and seem to take pride in their teaching, and it appears that any meddling from the school with either the creative process or the actual teaching method in class would be viewed as a negative obstruction.

Curriculum and class design.

Similar to the relationship between autonomy and teaching style discussed above is the relationship between autonomy and curriculum design. Akira, a tenured professor at a private university stated:

Actually, the dean of my college is relatively young. I think in his forties or something. He's really open-minded and actually encourages the teachers to get together and re-evaluate the English curriculum on our own. Make adjustments where we feel necessary, you know, keep it up to date with the research. It's often just the 4 or 5 of us English teachers getting together and putting the curriculum together. Nobody is looking over our shoulders. It's definitely a lot of work but exciting.

Likewise, Matt, an Australian contract teacher working at a private university made a similar comment regarding course design:

I teach this course 'Media English.' Actually, several of us teachers teach this same course. We get to design everything about that course from soup to nuts. Really it's our baby. Not easy, for sure, but it's great to have that kind freedom and responsibility, too. Nobody is shoving a pre-set curriculum down our throats.

Much like the universities where Matt and Akira are employed, Kate, a tenured professor at a private university, also remarked that one of the most positive aspects of her institution was actually being entrusted to develop a sound English curriculum with the other foreign English teachers. She remarked:

I feel lucky in a sense. In our department we have almost 30 full-time tenured faculty and 10 of them are native speakers of English. That's a huge dynamic, you see. There has always been a strong foreign presence and the school has always deferred to us for the English curriculum. Partly because of the school's support and their hands-off policy and partly because of the hard work of the native speakers, we have always been a curriculum innovator. Other nearby schools literally have copied our curriculum. When I oversaw the editing of the curriculum implemented a few years ago, we won the Good Practice award from the Ministry of Education. We did that!

What is noticeable from the three previous quotations is not only the obvious dedication the teachers have towards course and curriculum design, but the positivity, passion, and pride that is woven in to and conveyed through their comments, almost as a reward for being given the opportunity to take on a task, see it through, and succeed. Furthermore,

all three teachers seem to accept that the work will be difficult, but undeterred and even motivated perhaps, relish in being given an important responsibility.

Decision making.

Lastly, similar remarks were made concerning autonomy and general decision-making at Japanese universities. Kohei, a tenured professor at a private university noted the following:

Our university has many many rules and many many committees. Too many, actually. But what is better here than my last university is that the decisions are made after the teachers have a discussion. It's a longer process but it's superior. It's like a democracy. At my previous university the board of directors made all of the important decisions and we were told about them later. Here is a bottom-up system. The decisions are made by the teachers not the board.

Likewise, Will, an American tenured associate professor at a private university spoke about how the other teachers at his school not only wanted to hear his opinion, but also needed him when making decisions concerning the entrance examination.

I'm on the entrance exam committee this year. It sucks. I know. But I'm on the inner most part of the proverbial 'onion' so to speak. I'm pretty sure they trust me and the other committee members honestly want to know what I think about types of questions, style, and design of the test. They also kind of need me for the English, but it's still cool though. They're looking to me to make some of the final decisions. I've got to give the final ok before we submit the test.

In the cases of both Kohei and Will, it is clear that being included in the decision making process is viewed as a positive influence. Being on the inside and having a voice that is heard and acknowledged, certainly appears to have an encouraging effect on these teachers.

One central commonality from all of the above quotations concerning autonomy is the apparent hands-off "we trust you" approach taken by the institutions and the

various teachers' superiors. Giving the educators opportunities to succeed, giving them the freedom to make their own choices and decisions, and creating a positive atmosphere seems to be a key element in their appraisal of their own job satisfaction. Furthermore, these comments are in line with Bandura's (1977) research on the fourth source of self-efficacy: Physiological States, where he explains that mood affects people's judgments of their personal efficacy. Bandura stated that a positive mood enhances perceived self-efficacy and a despondent mood diminishes it. He goes on to explain that the fourth way of modifying self-efficacy beliefs is to reduce people's stress reactions and alter their negative emotional tendencies and misinterpretations of their physical states. Essentially, the creation of a good mood or atmosphere in the above teachers' schools and enabling their autonomy by trusting them and providing them with choices has the potential to support their self-efficacy beliefs.

Colleagues

Another positive theme that emerged frequently in the interviews was that of working with good supportive colleagues. The teachers reported that their colleagues helped motivate them in various ways, for example, by sharing knowledge, classroom management and teaching strategies, and lessons plans. They also stated that their cooperative colleagues made going to work easier and more pleasant.

Knowledge sharing.

In the case of Michelle, an American contract teacher, she professed about the readiness of her colleagues to share their knowledge of technology with her.

I'm terrible with technology because I'm afraid it will fail me. However, my colleagues are incredibly helpful because they are better with technology than me. They teach me all sorts of things. It's essentially working with these kinds of colleagues who are better than me that is making me a better more confident teacher.

Similarly, Yuka, a tenured professor, praised her colleagues for their willingness to share not only 'how-to' knowledge, but also lesson plans.

We learn by experience, right? But when you have very little experience, then what? My sempai, were good because they helped me when I first arrived. They showed me the communal cabinet in the office where the lessons they had developed themselves were kept. There was so much stuff I could use. A total lifesaver. They shared their ideas with me then and they still do now.

Matt, an Australian contract teacher, made a comparable statement about the abundant generosity of his colleagues when he first arrived at his university.

When I first arrived at the job, I got support. I knew I was fortunate because I had heard of other teachers at other schools being totally abandoned. But the teachers and I established strong relationships in the first few months. I love going to work. It's not unpleasant at all. They are super good to me and I like the people I'm working with now.

The three quotations above are analogous in that they all emphasize receiving help from colleagues when there was a lack of experience or knowledge. All of them were assisted at critical points in their teaching careers, leaving them better off and as Michelle said a "more confident teacher."

In my experience, there are many teachers at Japanese universities that have positive attitudes towards helping colleagues and often their intentions tend to be good regarding sharing of knowledge, too. Obviously, there are exceptions and teachers who don't actively go out of their way to help others or share what they know. However, in the cases of the three teachers above, it certainly seems that they are in very positive work situations and because of the generosity of their

colleagues, have improved and extended their relationships to those with whom they work.

Cooperation.

Akira, a tenured professor made two separate, but likewise comments about the cooperative attitudes of his colleagues. The first comment was:

I don't really have any complaints. The number of teachers in this department is quite small. 29 full time faculty, I think. They're all really cooperative. I haven't had any problems with them and the dean is not really imposing either.

And in a separate comment later in the interview, he said:

It's really been great. It's a new department and the faculty members are so cooperative. If I have a question, I can share it with people here. That's the best. And we always get together and decide the details together. It's very interesting.

There are three important points regarding the development and support of teachers' self-efficacy beliefs that can be interpreted from the encouraging comments above. The first point is similar to the earlier discussion about autonomy, where being given an appropriate amount of autonomy appeared to create a positive working environment that had the potential to support self-efficacy beliefs. Bandura (1997) explained that one way to raise self-efficacy beliefs is to improve physical and emotional well-being and reduce negative emotional states. Because individuals have the capability to alter their own thinking and feeling, enhanced self-efficacy beliefs can, in turn, powerfully influence the physiological states themselves. Therefore, it is perfectly logical that good cooperative colleagues could create such a positive working environment that it could in turn influence the physiological states of the teachers, which in turn could support the development of their self-efficacy.

The second point deals with the creating and strengthening of self-efficacy beliefs, which is through the vicarious experiences provided by social models. This is Bandura's (1997) second source of self-efficacy. At universities in Japan, the teachers interviewed in this study, seeing people similar to them succeed by sustained effort raises their beliefs that they too might possess the capabilities to master comparable activities required to succeed. By the same token, observation of other teachers failing despite high effort might lower their judgments of their own efficacy and undermine their efforts. The impact of modeling, sharing knowledge, and being helped as mentioned by the participants in their quotations above, on perceived self-efficacy is strongly influenced by the perceived similarity to one's peers. The greater the assumed similarity the more persuasive is their peers' successes and failures.

The last point concerns the third source of self-efficacy, Social Persuasions. Bandura (1997) described that individuals create and develop self-efficacy beliefs as a result of the social persuasions they receive from others. He stated that persuaders, for example, play an important part in the development of an individual's self-beliefs and that effective persuaders must cultivate people's beliefs in their capabilities, while at the same time ensuring that the envisioned success is attainable. Although none of the participants themselves made any explicit references to encouraging feedback (i.e., social persuasion) from their colleagues, I believe one can relate this source of self-efficacy to the comments from the teachers above, and it is fairly easy to imagine that the positivity they feel towards their colleagues because of the sharing of knowledge and lesson plans, generosity, and cooperativeness has created a warm working environment. The supportive words of encouragement and helpful actions the teachers received from their

colleagues, undoubtedly have the potential to support the self-efficacy beliefs of these teachers.

Money

Among Japanese private universities, salaries can vary greatly from school to school depending on such factors as their popularity, size, prestige, and location. Within each school, the salaries naturally fluctuate between ranks (i.e., assistant professor, associate professor, and professor), usually based on the school's own pay scale. At national universities however, this is not the case. There is a pay scale set by the government that is constant for all national university employees regardless the prestige, the location, or the size of the university. Although exceptions do exist, in general, private universities in Japan tend to pay significantly higher salaries than national universities. To put it in to an international perspective, compared to university salary averages in the United States, Japan's national university pay scale is relatively similar. For instance, according to the U.S. Department of Labor, in 2010 at a \$1 to ¥100 ratio, average salaries for full-time professors was \$98,974 in the US compared to a lower \$84,660 in Japan, \$69,911 for associate professors in the US and \$70,009 in Japan, and \$58,662 for assistant professors in the US compared to a higher \$62,223 in Japan (Bureau of Labor Statistics, 2012). Finally, it should be noted that in addition to the Japanese base salaries, there are often quite a few supplementary and sizeable allowances added to their monthly salaries. These allowances often include money for research, dependents, housing, transportation, and managerial work, substantially increasing the overall

income. These allowances, should they have been included to the numbers above, would undoubtedly push the average Japanese salaries much higher.

Money (e.g., salary, bonuses, allowances), which is an extrinsic motivator, on its own is not a direct source of self-efficacy. However, appropriate compensation can lead to positive feelings for teachers, thus supporting the creation of high self-efficacy beliefs. The connection with self-efficacy is that if the teachers perceive adequate compensation from their universities and they believe that they have the required resources and opportunities to excel, and that the obstacles they are likely to encounter are few and manageable, they should have more confidence in their ability to teach and thus exhibit a high degree of self-efficacy. Conversely, when they believe that they lack requisite resources or that they are likely to encounter serious financial obstacles, they are susceptible to judge their teaching to be relatively difficult and hold a low level of self-efficacy.

Positive extrinsic motivators such as salary, pension, insurance, and research benefits, have long been known to contribute to job satisfaction and teacher motivation. It is therefore not surprising that many of the interview participants in this study also remarked on how the various types of monetary compensation that the universities awarded them, were viewed very favorably. Some teachers commented on the generous amounts of money allotted to support their research, while others commented directly on the high university salaries. In either case, it became quite clear that being paid adequately and appropriately definitely mattered to these teachers.

Research.

Kenta, a tenured professor at a private institution in western Japan, commented on a few of the research-related positive financial opportunities at his university.

As you can see (simultaneously showing me his new iPad, iPod, MacBook air, and other such devices), I enjoy a comfortable research environment...the money we get for research is really good. I make several overseas trips a year to go to conferences with that money.

Kohei, also a tenured professor, made similar comments about his more than adequate research budget.

They are totally supportive for research grants. They will also pay for each professor to make two presentations at international conferences as well as at three or four domestic conferences. On top of that we get nearly 500,000 yen in kenkyuhi (research allowance). And if you need it, you can apply for extra kenkyuhi to double your amount. I don't know where they get the money, but they are generous.

Will, an American associate professor at a private university, also gave his frank opinions about the merit-based allotment of research funds at his school.

The research money you get here is ridiculously high but also quite competitive. It all depends on what you produce. The first year you arrive, you get a bunch of money. If you would like to receive that same amount of money the next year, you need to produce. You need to publish or present or be on a research team or in the middle of a research project or something. They are pretty reasonable though. The requirements are not too strict and the people in charge seem fair for the most part. It keeps people motivated and as far as I can tell, everyone seems content.

University teachers in Japan commonly receive a generous research budget at the beginning of every academic year. Many have come to take it for granted. The amount varies from school to school but is usually in the range of ¥200,000 to ¥400,000.

Additional funding might also be available to the faculty on a competitive basis either through the school, local private companies, or government grants. This is in sharp

contrast to the United States, where nearly all additional funding for research is received after a strict application process and is entirely competitive and merit-based. While the teachers spoke to me in very matter-of-fact terms about these funds, what is strikingly clear and common about the three quotations above is that the teachers seem more than pleased and satisfied with their large research allowances.

Salary.

Similar to the research allowances, the teachers also spoke positively about their salaries, too. Kenta, a tenured professor at a private institution in western Japan, commented enthusiastically about his salary when he said:

My salary is not so bad either (smiling). In general we are really well paid. We get extra allowances for everything, like teaching night classes or working on entrance exams, etc...

Robert, an American associate professor at a national university in fact expressed some feelings of guilt about the money he receives.

People complain too much. Financially speaking, I almost feel guilty. They compensate us very well. And really, for the amount of classes I actually teach and the amount of time off we have in between semesters, how can I complain?

Finally, Matt, a contract teacher at a private university, painted a very positive and clear overarching picture about the benefits of the money and lifestyle afforded to him by his university and living in Japan. This quotation was taken from a follow up e-mail after the interview.

The industry is generally well developed and stable in Japan, which offers an economy and lifestyle that affords opportunities to do more things and go more places and reach other people if a person so desires. Or to simply pad one's nest egg or mattress as this is as effective as putting the money earned, which can be substantial compared to two classes a week at a Community College, if they even

get enough people, into the bank. As someone who works hard and plays hard, the financial opportunities at my university have satisfied me on many levels, personally, professionally, socially, mentally, and physically.

Appropriate financial rewards for teachers are an extremely sensitive, yet also very important issue when discussing job satisfaction and teacher self-efficacy. The importance of remuneration is summed up well in this quote from Poppleton and Riseborough (1990),

Pay does not have absolute importance in relation to job satisfaction but if it is perceived to be good all other aspects appear to have relatively less significance. If, on the other hand, it is perceived to be poor, then it is seen as a symptom as much as a cause and associated with other symptoms such as lack of respect in the community. (p. 219)

Self-efficacy research tells us that teachers with perceptions of greater self-efficacy are more likely to have higher performance standards and goals, have more favorable job attitudes, and show greater willingness to put forth effort on challenging tasks. Therefore, appropriate compensation is an important factor that has the potential to support language teacher self-efficacy.

Students

The final positive theme likely to support teacher self-efficacy that emerged frequently in the interviews was that of teaching good hardworking students. Not surprisingly, many of the interviewees commented that having good students made coming to work everyday enjoyable. They often elaborated stating that good students were the best part of their jobs and if it were not for such students, they might consider changing professions or at least consider changing universities.

Naturally, not all of the participants are able to teach super hardworking motivated students. Additionally, there is always going to be a lot of variation depending on such factors as where the faculty are teaching, which populations they are teaching within their university, and how many of the courses they teach are electives vs. required. In the analysis below, I will try to simultaneously explain the various teaching situations of the participants in order to help contextualize their comments. The following quotations highlight the positive points the interviewees spoke about their students.

Michelle, a non-tenured American teacher working at a private university described her feelings about her students as follows:

The best part of my job is working with all of these motivated students. They can make miracles happen in the classroom. I love that. I am very happy with my students and wouldn't want to teach anywhere else.

Michelle works with above average well-rounded students. Many of them come from well-to-do upper middle class families and as she describes them, "they are just good kids who obviously come from good families who raised them right." She explained that behavioral problems with these students were extremely rare and although she mainly teaches compulsory four-skills courses, the classes always have fewer than 20 students making classroom management almost a non-issue.

Robert, an American associate professor at a national university expressed similar sentiments when he said:

The easiest and perhaps best part of my job is the students. Period. Maybe it's because that's where I focus my effort or maybe it's because they are just really good students. It's a chicken and the egg situation. Either way makes going to work quite gratifying.

Robert works at a large and prestigious university in eastern Japan. The student body as a whole is quite capable and certainly above the national average. As he puts it “at the very least, they studied hard enough in high school and learned to pass a rather difficult entrance exam. They might not have the English skills yet, but they have the work-ethic.” He also explained that his classes are generally small and manageable with the odd class perhaps having a few more than 20 students.

Kohei, a tenured professor at a private university also expressed the importance of having good students.

The most important thing to me is having good students or at least students who will try their best. Here, it’s so easy to have a friendly chat with motivated students before, during, or after class. The kind of students that you can talk about anything with and that are eager to learn... It’s what makes this job worthwhile for me.

Similar to Michelle and Robert’s teaching situation, Kohei also teaches mid to upper level students. He works at a well-known wealthy private school with many students entering from strong Tokyo-based high schools. Teaching not only compulsory classes, but also upper-level elective seminars, Kohei has the advantage of working with students who not only are motivated and have chosen to take his class, but also small numbers of students in those classes.

Finally, Kate, an American tenured professor at a private university, summed up her feeling about good students as such:

The best part of my job is the students. For the most part they are capable and very pleasant. If it weren’t for them I’m not sure I would have been here this long. Even the low level learners, they are just good kids who try. Even if they don’t have the ability to do it all in English, they’ve got brains and at least take a shot at using them, which is more than I can say for some of my colleagues. I think this is our students’ biggest strength. We provide the possibility and they put forth the effort. They really do and I love it.

Kate's working situation is almost identical to that of Michelle's in terms of the type of students (e.g., upper class families, raised right, few behavioral issues). Furthermore, she also has a lot in common with Kohei's teaching situation in that being a veteran tenured teacher means she gets to teach seminars to small groups of students who have selected her classes.

Two of the most noticeable commonalities in the teaching situations described above were the small class sizes and the well-rounded hardworking students. All of the teachers work in good schools and despite inevitable variations in the students' English levels, they all seem to have an understanding of how to study hard. Additionally, the small class sizes, which mean more individualized attention, can only help lead already good students to excel even more.

The relationship between good students and potentially increased teacher self-efficacy and job satisfaction is an easy one to understand. As Lortie (1975) remarked, teacher-student interaction was a major source of overall satisfaction and claimed that:

Other sources of satisfaction pale in comparison with teachers' exchanges with students, and the feeling that students have learned. We would therefore expect that much of a teacher's work of motivation would rotate around the conduct of daily tasks - the actual instruction of students. (p. 104)

Furthermore, Skinner and Belmont (1993) found that teachers' perceptions of student emotional and behavioral engagement, in fact predicted teachers' interactions with their students. They found strong support for positive student engagement eliciting positive teacher behaviors. Moreover, Skinner and Belmont's research demonstrated that teachers respond to students who have high behavioral engagement "with more involvement, more autonomy support, and even to a degree, more contingency and consistency" (p. 578).

Additionally, Bandura (1997) stated that people's beliefs about their efficacy are developed by four main sources of influence. The most effective and influential source is through hands-on mastery experiences. For a teacher, these experiences will undoubtedly occur in the classroom. Successes in the classroom will build a robust belief in one's personal efficacy. Teachers will engage in tasks and activities, interpret the results of their actions, use the interpretations to develop beliefs about their capability to engage in subsequent tasks or activities, and act in accordance with the beliefs created. Typically, outcomes interpreted as successful raise self-efficacy. Therefore, it is entirely logical that having good students that are well behaved, eager to learn, and easy to engage, at the very least, will create an atmosphere where self-efficacy beliefs can flourish.

Research Question 6: Potential Self-Efficacy Challenges

Administration

Many Japanese and foreign teachers pointed out, often with great disdain and contempt, that some of the most negative aspects of their jobs were the various difficulties that arose with their direct superiors, the general administration, or both. They particularly mentioned that problems such as lack of communication and poor management by the administration, had made their professional lives extremely trying and problematic, and at times entirely unbearable. Before examining the interview data, the following briefly describes selected research on the topic of faculty and administration relationships.

In a study by Ehara (2000) examining faculty perceptions of university governance in Japan and the United States, he found that just under half of Japanese and

U.S. faculty overall (46.2 percent) believed top-level administrators provided competent leadership, while 52.3 percent believed the administration is often autocratic. He also found that many faculty members desired improved communication with administrators, with only 38 percent of faculty overall indicating that they are kept informed about what is going on at their institution. Overall, 44 percent believed communication between the faculty and administration is poor. This dissatisfaction is highest in Japanese private universities but is not regarded as much of a problem in Japan's public universities.

With the exception of faculty in Japanese public universities (20 percent), more than 40 percent of faculty considered their lack of involvement in decision making a serious problem. However, it is not necessarily the case that creating more opportunities for faculty participation would increase faculty satisfaction. For example, while faculty considered expanded participation important for the university's management and administration, there is also resentment that too much of their time is spent in this way.

In line with the research cited above, a few choice quotations that give an overarching picture of the struggles experienced by the interviewees with their respective administrations are presented below.

Lack of communication.

First, Kenta, a tenured professor at a private university vented his frustrations regarding lack of communication and university politics when he said:

Politics! I'm tired of the so-called sectionalism. The administration is overly conservative and they have all these trivial distinctions at the expense of the general well being of the teachers. Nobody communicates either. There's not enough talk between teachers and the administration, and therefore nobody knows

what's going on. You always have to fight to get something changed. It's really exhausting!

Similarly, Kate, an American tenured professor at a private university expressed her grievances concerning the lack of communication between the administration and the teachers when she said:

Communication with the administration has been very poor for many years. I know it sounds like I'm just complaining, but foreigners and Japanese alike are so frustrated with them. There are too many people here who work *in* an education system that don't know how to work *with* an education system. They do not communicate and cannot understand being in a classroom and the higher-ups especially, don't understand that you can't tell a teacher to do things in the same way that you can order a salary man to do things in a company. Sometimes I honestly think they haven't a clue how to run a university effectively.

The two quotations above share the palpable frustration felt by the teachers created by the lack of dialogue between those who hold administrative roles and those who teach. There seems to be a great chasm between the two groups over what each group believes is best. Nonexistent communication only widens the gap and in Kate's situation in particular, she eludes to the fact that perhaps many people working in the administration do not understand what it means to work with an education system and instead simply treat it as any other regular company, an idea with which she strongly believes does not work.

Poor management.

In a follow-up e-mail, Robert, an American contract associate professor at a national university expressed sharp dissatisfaction about the poor management of the language program and his superiors not being invested in the program when he said:

When it comes to the language program here, there seems to be a general lack of accountability and responsibility among my superiors. It's become an administratively-focused approach to curriculum and program operation, with the

academics who should have a stake in the results of the program being too hands off, on a rotating committee, and generally being out of touch and not vested in the program or the process and its outcomes. Fat cats riding on a fat cow!

Likewise, Matt, an Australian contract teacher at a private university, paints a clear overarching picture about the difficulties he faces with the administration at his university.

Addressing the range of individual needs and personalities at the human level outside the classroom with various stakeholders and their petty privileges and fiefdoms has been and will always be my scourge. The useless make work projects and tacit lip service paid to certain aspects of the 'big picture' end up taking more of my time and focus away from what we are supposed to be doing, namely teaching and researching. Those other things are 'someone else's job' and if we lacked in our job to the same degree as the administration it would be very visible and awful and the hypocrisy that exists is laughable.

The gripes that the interviewees had with their respective institutions' administration were conveyed vividly. Additionally, the lack of institutional supports and lack of communication that persists between the teachers and the administration, obviously play a vital role when looking at the importance of self-efficacy. For example, Mowday and Nam (1997) stated "people are more likely to engage in behaviors when they see a high probability that effort will lead to high performance" (p. 117). Connected to the idea of teacher self-efficacy, the more support teachers receive from their institutions and the better the relationships are with all involved parties, the more likely teachers will feel their efforts will be successful, and are thereby more willing to put time and energy into their jobs.

Workplace environments absolutely influence self-efficacy beliefs, with supervisors' verbal persuasion and modeling serving as important spurs to workers' self-efficacy development (Bandura, 1997). Kooij et al. (2008) suggested that age-related

physical and psychological factors can influence work motivation, but workplace factors can mediate how age-related concerns are interpreted. For teachers, the combination of successful past experiences, verbal support from students, peers, and parents, as well as supervisors, and opportunities for observation of successful peers builds self-efficacy for teaching (Tschannen-Moran et al., 1998). The influence of the sources of self-efficacy, however, can change over time, with verbal persuasion and contextual factors playing a more important role for novice teachers than for veteran teachers (Tschannen-Moran & Woolfolk Hoy, 2007). Self-efficacy beliefs in the workplace are not static and reflect a lifelong process of development that ebb and flow according to personal attributes and interpretation of environmental circumstances (Klassen, 2010).

If better support is established and if teachers believe that their efforts will lead to success and not be hindered by perceived incompetence of the administration, they will likely be more motivated and work harder, and as Matt's quotation eludes to, spending more time for example, on preparing for classes, being more available to students, and researching.

Furthermore, Olsen (1993) demonstrated through her survey of first and third year university faculty that many teachers felt that better relationships with the administration would positively influence their teaching. Likewise, Freeman and Philips (1985) also indicated that many teachers blame too little institutional support and too many restrictions for their lack of successful classroom outcomes, and thereby lose intrinsic and extrinsic motivation due to a lack of self-efficacy beliefs. In short, if teachers do not believe they are getting adequate institutional support, they will less likely believe in the possibility of success, hence severely lowering their perceived self-efficacy.

Students

Research shows us that students are one of the main sources of motivation and demotivation of teachers. Vandenberghe and Huberman (1999) wrote, “The quality of the relationship between teacher and pupils can be one of the most rewarding aspects of the teaching profession, but it can also be the source of emotionally draining and discouraging experiences” (pp. 194-195). It is therefore not surprising that many of the interview participants in this study remarked on not only how wonderful the students were, but also how grueling they could make their life at work.

As in the previous section about the positivity of students, there is always going to be a lot of variation depending on such factors as where the faculty are teaching, which populations they are teaching within their university, and how many of the courses they teach are electives vs. required. In the analysis below, I will again try to explain the various teaching situations of the participants in order to help contextualize their comments. The following quotations from the interviewees touch upon the obstacles and barriers they face with their students.

Akira, a tenured professor at a private university commented on how he actually left his previous position at a nearby institution due to the overwhelmingly unmotivated student body.

I was teaching English at an engineering college. It was really tough. They were not motivated at all. I tried so hard to emphasize that English was important for their futures, especially because they were engineering students, but they never listened. They could not see the connection and although I was tenured, I was there for just three years before I quit. I couldn't teach such poor students.

Akira is referring to a low level private engineering school in central Japan where he used to teach. His class sizes were large (usually 40 to 50 students) and the classes he taught

were mainly compulsory English courses. Many of the students in the school went to local low-level technical high schools and according to Akira had “little to no motivation.”

Masato, an experienced contract teacher at a private university shared his views on the downward trending level of Japanese university students and the difficulties it is causing.

I’ve spent time in the private sector and therefore tend to put great demands on my students and expect a lot. My goal is simple: to raise their English ability enough so they can be competent in the industry. But historically in Japan, university is paradise and the students don’t work hard. There are many complaints recently (from the companies) that Japanese students are no longer competitive and don’t have guts. The university regulations have become so loose and it’s just too difficult to get through to these students.

Although Masato now teaches at a well-known “elite” private university in Tokyo, he has over 25 years of teaching experience at various Japanese universities as well as over 10 years working in the private sector, and is in a prime position to comment on general trends at Japanese universities. It’s interesting that he feels the decline in ability has not only affected the lower ranked schools but the higher tiered schools as well.

Also in a good position to comment on general academic trends, Kentaro, a professor with over 20 years of experience at an upper-tier private university, made discouraging comments indicating his frustrations about how the previous generations of students were much better compared to those of today.

My biggest frustration is the students. I don’t think it is isolated to just this school either. The student population all over Japan has changed for the worse compared to just 20 years ago. I mean that the proficiency levels and motivational levels were so much higher back then. Many students now are not willing or don’t actually know how to study. They just come to university because they don’t want to work or their parents told them to go. They’re not paying anyway and are here for the wrong reasons! How am I supposed to deal with that?

Similar to Masato, Kentaro's quotation again shows us that the general decline in the students' ability is not limited to the lower level universities.

Finally, Rie, a contract instructor at a private university, summarized her vexations with her university English students.

Babysitting and scolding my students is what I like least of my job. They don't think they will need English for their futures and motivating them is the hardest part. They have an apathetic attitude toward learning and they use Japanese excessively in class. I regularly fail miserably at keeping their attention and it absolutely is shattering my confidence. At the worst times, I don't even want to come in and teach.

Unfortunately for Rie, she teaches exclusively compulsory English classes at a mid-level university to first and second year students. Because she is stationed at the language center, she teaches students from a wide variety of departments with greatly varying motivation levels. While occasionally she gets to teach students from one of the "good" departments, more often than not she explains that her classes are with students who would prefer not studying English at all.

According to a national survey in the United States of 1,920 faculty members by Gmelch, Wilke, and Lovrich (1986), students were one of the faculty's top five sources of stress. Similarly, Gates (2000) reported a study (Sax, 1996) where 61% of faculty felt that students were a major source of tension and anxiety. Furthermore, he concluded that dispirited or unprepared students or students who attended the university for the wrong reasons were among the major sources of faculty angst. Likewise, Guskey and Passaro (1994) reported that if teachers encounter obstacles they cannot deal with, they would become demotivated. Walker and Symons (in Bess, 1997), in discussing social motivation theory, report that if teachers perceive negative feelings on the part of the

students, and if they feel unappreciated in spite of their best efforts, then they will respond with anger and anxiety. Finally, Sergiovanni (1967) concluded that a poor relationship with students could be a source of considerable teacher dissatisfaction.

In addition to the research described above and as I mentioned in the introduction of this dissertation, one of the many reasons for the declining academic level of student in Japan, is the demographic trend of declining birthrates that has created a situation where the number of college-aged young people has decreased by so much that the country has reached a period of full college and university admissions. This means that any Japanese person who wants to go to college will have a 'seat' somewhere at a university in the country. One of the most obvious repercussions is the predicament of how to meet admission quotas and face the increasing lack of competition for admissions to many institutions. By eliminating competition and offering a place to all applicants, a general decline in the overall academic abilities of incoming college students is inevitable. Institutions quite simply have had to lower their standards and admit students that would not have been accepted even a decade ago (Hani, 2001).

Teaching can bring personal satisfaction, but it also brings stress with real demands from various sources, especially students. This stress can be substantially compounded with student misbehavior and lack of student motivation (Greenglass & Burke, 2003). Teachers with greater teacher stress, defined as "the experience of negative emotions resulting from a teacher's work" (Kyriacou, 2001, p. 27), have lower self-efficacy (Schwarzer & Hallum, 2008; Skaalvik & Skaalvik, 2007), poorer teacher-pupil rapport, and lower levels of effectiveness (Kokkinos, 2007). Furthermore, teachers with high levels of job stress can still gain satisfaction from work, but the level of satisfaction

might be greatly subdued by stress from the frequency or level of conflict with their students (Greenglass & Burke, 2003).

Finally, the already delicate reciprocal relationship that exists between students and teachers as described in the previous research above, combined with the declining birthrate and downward trending academic levels in Japan, has created a perfect storm for teachers at Japanese universities. It is these very teachers who are expected to respond to the demands of the parents, the universities, and government to raise the level of English of the now less-able and less motivated students. Dealing with these kinds of learners in the current climate of Japan could easily challenge a teacher's self-efficacy.

Contract Teachers

The final category likely to challenge teacher self-efficacy that emerged frequently in the interviews was that of limited-term contracts and the adverse affect they had on the faculty, the curriculum, the students, and the institution as a whole. Not surprisingly, many of the interviewees could find no merit whatsoever in such limited-term contracts and commented that they created unwanted stress, less than ideal working conditions, and by their very nature were inherently demotivating. The following quotations highlight the troubled feelings the interviewees had about limited-term contracts for English teachers at Japanese universities.

Rie, a contract English teacher at a private university, expressed her feelings about the chaos that is created from teachers always coming and going.

In the current 'fukeiki' (declining economic environment), teachers will continue to leave their contracts early because of the competitive job situation. The Foreign Language Center provides no light at the end of the tunnel and certainly no

opportunity for tenure and therefore will continue to be a revolving door of English teachers. This university will use the teachers and the teachers will use the university's name to get their next position. Teachers will continue to leave at inopportune times causing confusion, a lack of continuity, and a lot of extra work and headaches for the administrators and other teachers, too.

Similarly, Robert, a contract associate professor at a national university, pointed out the bleakness that is felt from the first day of a contract position is entirely uninspiring. He said:

Knowing from day one that we are limited to a set number of years, new English teachers will never be invested enough in the future of this university's English program. They have essentially guaranteed us unemployment. The teachers know that they are disposable and that there is absolutely no chance to see the fruits of their labor...Talk about demotivating!

There is a strong commonality in the two previous quotations in that both Rie and Robert address the issue of "no opportunity" and "no chance" whatsoever to obtain a tenured status and therefore making them feel completely "disposable". Undoubtedly for these two teachers, the lack "light at the end of the tunnel" is at best uninspiring and at worst, downright depressing.

Masato, an experienced contract teacher at a private university as well as having spent many years working in the private sector, expressed levelheadedly his candid feelings about the shortcomings of limited-term contracts.

These contracts reflect general industry, downsizing, outsourcing, etc... Honestly, it makes sense from an economic standpoint but doesn't mean it's good for education. It's actually creating a great deal of instability. This is unfortunate but very difficult to change. Term-limited contracts will have more of a negative impact and affect the quality of teaching because these teachers will always be looking for their next job and can't be devoted as much as they should be to the courses at their own university.

When I spoke with Masato, I sensed a feeling of helplessness from his inability to change anything about the limited-term contracts situation. At 65 years old and nearing

retirement, although he is not personally affected by such limited-term contracts, having spent a huge portion of his life in academia, he is noticeably frustrated with the current direction Japanese universities are heading.

Finally, Will, an American tenured associate professor at a private university spoke about the limited-term contracts versus the tenured positions from a frustrated foreigners point of view.

I cannot think of any benefit of being a contract foreign teacher. Sure, you get to skip out on boring meetings but c'mon, work is work. This short-term contract system for foreigners was established a long time ago and is no longer relevant. It's 2012 and many of the foreigners in Japan are no longer visitors. Many of us have families and people who depend on us. This job is not a joke, not a gig, and we are not just hanging out in Japan. We have the same qualities as the Japanese staff, we are professionals, we pursue higher degrees, we research, we teach, we maintain administration duties, and many of us have very high levels of Japanese proficiency. So why the distinction?

In recent years, various critics within Japan have extensively examined the conditions surrounding limited-term contracts for university language teachers. Furthermore, the connection between “guaranteed unemployment” as one of the interviewees suggested, and the negative influence it could have on one’s self-efficacy is unmistakably clear. For example, some of the greatest apprehensions for many teachers have to do with employment instability and the anxiety generated from the continual cycles of changing employment, as well as the psychological, physical, and monetary hardships of periodically moving the family unit. Even for those without family connections, Rivers (2013), claims that the nomadic lifestyle that limited-term contract employees tend to live, often inhibits the formation of sustainable collegial relationships, restricts workplace involvement in long-term initiatives, denies emotional attachment to a specific place (i.e., developing a sense of home or belonging), and undermines sincere

dedication to one's contracting institution. Every point that Rivers mentions above, as well as the almost obsessive like quest to continually search for improved working conditions that many contract teachers undergo, simply does nothing to promote self-efficacy.

Additionally, limited-term contract teachers' work in Japanese universities is characterized by heavier teaching loads, insecurity caused by contract status, little input into or control over teaching assignments and curriculum, lack of time for research, and their consequent devaluation as "teaching-only" faculty. Furthermore, Bandura (1997) explained that the teacher's beliefs in their coping capabilities affect how much stress and depression they experience in threatening or difficult situations, as well as their level of motivation. He went on to explain that perceived self-efficacy to exercise control over stressors (i.e., lack of job security) plays a central role in anxiety arousal. People who believe they can exercise control over such threats do not conjure up disturbing thought patterns. But those who believe they cannot manage threats experience high anxiety arousal. The teachers will ultimately dwell on their coping deficiencies. They might view many aspects of their teaching environment as troubled and in jeopardy of being taken away. Additionally, the teachers in such environments will very often magnify the severity of possible threats and worry excessively. Through such inefficacious thinking they will distress themselves and impair their level of functioning. The teacher's perceived self-efficacy regulates avoidance behavior as well as anxiety arousal and the diminished sense of self-efficacy the less likely teachers are in taking on more taxing and threatening activities.

As listed above, in dealing with what appears to be a mountain of impediments, it can safely be said that limited term contracts for English teachers at Japanese universities, certainly pose a threat to the teacher's self-efficacy.

Summary

In the present study, the interviews were a complementary data source used for answering Research Questions 5 and 6. Thus far, I have presented the results and a discussion of the interview data analysis according to each of the seven emergent themes. Four themes found in the data (Autonomy, Colleagues, Money, and Students) helped answer Research Question 5 and spoke to qualities that could support teachers' self-efficacy. Three themes found in the data (Administration, Students, and Limited-term Contracts) helped answer Research Question 6 and spoke to qualities that could weaken teachers' self-efficacy. From this chapter, it is clear that the individual teacher's beliefs and assessments of their own working situations and institutions are not only multifaceted and unique, but also have many overlapping qualities, namely the aforementioned seven themes. Without knowing the intricacies of such university contexts and the complexities of the individual teachers' feelings, discussions on how to improve their self-efficacy beliefs will likely remain unsuccessful.

The interviews provided me with details about both the positive and negative influences in a Japanese university English teacher's work life. In the next chapter, I summarize and discuss the quantitative results.

CHAPTER 7

DISCUSSION

In Chapters 5 and 6, the quantitative results (i.e., questionnaire) and the qualitative results (i.e., interviews) were reported. In the case of Chapter 6, the results were not only reported but also examined and discussed. Because the present study adopted a triangulation strategy mixed-method design, the quantitative data were considered the main data source to answer the four primary research questions while the qualitative data served as a complementary data source to explain research questions five and six. Therefore, in this chapter when discussing the main results, the qualitative data that were examined in Chapter 6, only where relevant and according to the emergent themes discovered, are selectively referred to for the purpose of explaining, extending, and contextualizing the survey findings (i.e., research questions 1-4). Furthermore, the following discussion is developed around Bandura's (1986) social cognitive theory, reciprocal determinism, and self-efficacy construct. Finally, at the end of the chapter, contributions and implications are offered.

Social Cognitive Theory and Reciprocal Determinism

As discussed in the first two chapters of this study, Bandura (1986) advanced a view of human functioning that accords a central role to cognitive, vicarious, self-regulatory, and self-reflective processes in human adaptation and change. People are viewed as self-organizing, proactive, self-reflecting and self-regulating rather than as reactive organisms shaped and shepherded by environmental forces or driven by

concealed inner impulses (Pajares, 1996). Bandura altered the label of his theory from social learning to social “cognitive” both to distance it from prevalent social learning theories of the day and to emphasize that cognition plays a critical role in people’s capability to construct reality, self-regulate, encode information, and perform behaviors.

From this theoretical perspective, human functioning is viewed as the product of a dynamic interplay of personal, behavioral, and environmental influences. For example, how people interpret the results of their own behavior informs and alters their environments and the personal factors they possess which, in turn, inform and alter subsequent behavior (see Figure 1 in Chapter 1). This is the foundation of Bandura’s (1986) conception of reciprocal determinism, the view that (a) personal factors, (b) behavior, and (c) environmental influences create interactions that result in a triadic reciprocity. The reciprocal nature of the determinants of human functioning in social cognitive theory makes it possible for beneficial efforts to be directed at personal, environmental, or behavioral factors. Strategies for increasing well being can be aimed at improving emotional, cognitive, or motivational processes, increasing behavioral competencies, or altering the social conditions under which people live and work.

In schools, for example, Pajares (1996) explains that teachers have the challenge of improving students’ academic learning and confidence. Using social cognitive theory as a framework, teachers can work to improve their students' emotional states and to correct their faulty self-beliefs and habits of thinking (personal factors), improve their academic skills and self-regulatory practices (behavior), and alter the school and classroom structures that can work to undermine student success (environmental factors).

Finally, of all the thoughts that affect human functioning, and standing at the very core of social cognitive theory, are self-efficacy beliefs, “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura, 1997, p. 3). Self-efficacy beliefs provide the foundation for human motivation, well-being, and personal accomplishment. This is because unless people believe that their actions can produce the outcomes they desire, they have little incentive to act or to persevere in the face of difficulties. As seen through the review of the literature in Chapter 2, much empirical evidence now supports Bandura’s contention that self-efficacy beliefs touch virtually every aspect of people’s lives, whether they think productively, pessimistically or optimistically; how well they motivate themselves and persevere in the face of adversities; their vulnerability to stress and depression, and the life choices they make (Pajares, 1996).

Bandura’s (1997) main arguments as concerns the role of self-efficacy beliefs in human functioning are that “people’s level of motivation, affective states, and actions are based more on what they believe than on what is objectively true” (p. 2). For this reason, how people behave can often be better predicted by the beliefs they hold about their capabilities than by what they are actually capable of accomplishing, for these self-efficacy perceptions help determine what individuals do with the knowledge and skills they have. This helps explain why people’s behaviors are sometimes disjoined from their actual capabilities and why their behavior can differ widely even when they have similar knowledge and skills. For example, many talented people might suffer incapacitating bouts of self-doubt about capabilities they clearly possess, just as many individuals are confident about what they can accomplish despite possessing a humble inventory of

skills. Pajares (1996) explains that belief and reality are rarely perfectly matched, and individuals are typically guided by their beliefs when they engage the world. As a consequence, people's accomplishments are generally better predicted by their self-efficacy beliefs than by their previous attainments, knowledge, or skills. Of course, no amount of confidence or self-appreciation can produce success when requisite skills and knowledge are absent.

The subsequent discussion will explain the results of the current study, while simultaneously trying to situate them in the theories and ideas above.

Profile Analysis

As explained in depth in Chapter 5, profile analyses were chosen to answer the first four research questions. In the profile analyses performed in this study, the data were plotted with the self-efficacy factors on the x-axis and the response scores on the y-axis. These plots were then made into profiles or lines, representing the mean scores across dimensions (i.e., self-efficacy factors) for each group. In this study, the profile analyses asked these three basic questions about the data plots: (a) Are the groups parallel between self-efficacy factors? (b) Are the groups at equal levels across self-efficacy factors? (c) Do the profiles exhibit flatness across self-efficacy factors? If the answer to any of these questions was no (i.e., that the null hypothesis was rejected) then there was a significant effect. Finally, because much interpretation can be done using graphical methods, in the ensuing discussion about each profile analysis performed, the reader is encouraged to simultaneously refer to the four Figures presented in Chapter 5 for a visual representation of the results.

Research Question 1

The first research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' native language. This question was answered by conducting a profile analysis, where the teacher's native language (English or Japanese) was entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors obtained from the questionnaires were entered as the dependent variables.

The answer to the primary question, "do the native English teachers and Japanese English teachers in the study have parallel profiles," was no, indicating a significant effect. This test revealed that the native English teachers and Japanese English teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared. Although the same up-down trend of the lines for both native speakers and non-native speakers can be seen in Figure 12, the fact that the lines have marginally different slopes at different points denotes a significant effect. The answer to the second question, "are the two groups at equal levels across observations", was also no, indicating a significant effect. In other words, this test showed that one group scored higher on average across all measures. The results indicate that native English teachers, on average, have reliably higher self-efficacy scores than native Japanese teachers on the selected set of measures. This test is displayed in Figure 12 by the fact that the lines are not right on top of each other and that the four respective points on each line are separated by a statistically significant gap. The answer to the third question, "do the profiles exhibit flatness across observations", was also no, indicating a significant effect. In other words, the flatness test checks the similarity of responses to all dependent

variables, independent of groups, and in this context, the profiles did not exhibit flatness across observations. Because the result was significant, it suggests that teachers' self-efficacy varied regardless of whether the teacher was a native English speaker or native Japanese speaker. In Figure 12, this is shown by no two points on the same line to have the same score.

When looking at Figure 12, perhaps the first and most obvious feature is that every self-efficacy variable is noticeably higher for the native English speakers than for the Japanese speakers. This suggests that on average, the native English teachers perceived themselves more able regarding the engagement of their students, in their instructional strategies, in their classroom management techniques, and in dealing with their superiors.

Efficacy beliefs are typically “self”-beliefs: They consist of statements and thoughts about “I” and “me” and make up “what we bring to mind when we think about ourselves” (Neisser, 1997, p. 4). In a Western context, perceptions of efficacy are reported to influence perseverance, resilience, and task choice (Bandura, 1997). Various researchers (see Pajares & Graham, 1999) have found self-efficacy beliefs to be stronger predictors of subsequent performance than past performance or other motivation constructs. Nevertheless, a particular way of thinking about our selves might not be universally held: Markus and Kitayama (1991) propose that self-phenomena can assume a different form according to the culturally influenced relationship of self with others, as described by the cultural dimensions of individualism and collectivism. The cultural dimensions of individualism and collectivism refer to the degree of separateness and connectedness of individuals and groups (Triandis, 2001). Individualist cultures like the

United States, Canada, England, and Australia tend to emphasize “I” consciousness, independence, individual initiative, and right to privacy. In contrast, collectivist cultures like China, Japan, and Korea have a tendency to stress “we” consciousness, collective identity, group solidarity, and duty (Kim, 1994). Unfortunately, like many psychological theories that aspire to universal relevance, self-efficacy has been conceived and studied almost exclusively in Western settings, with little attention paid to the 70% of humans with non-Western cultural backgrounds (Triandis, 2001).

Citing relevant studies on cultural difference and self-efficacy, the following is a discussion of why, perhaps, the native English speakers rated themselves significantly higher than the Japanese teachers on the questionnaire. In a study examining the achievement and motivation of Asian American and non-Asian (predominantly Anglo-American) high school students, Eaton and Dembo (1997) found that the non-Asian students rated their self-efficacy, measured at the specific task level, significantly higher than did the Asian American students. However, performance on the task (unscrambling 20 words from a reading passage) was significantly lower for the non-Asian students. Results from a multiple regression analysis predicting achievement motivation showed that for the Asian Americans, fear of academic failure was the strongest predictor of achievement, followed by self-efficacy; for the non-Asians, fear of failure was not a significant predictor, while self-efficacy beliefs provided a significant contribution. The bivariate correlations, which can be seen as a measure of calibration between efficacy beliefs and performance, were stronger for the Asian students; i.e., in this study, Asian students were more accurate at calibrating their efficacy beliefs with subsequent

performance. The Asian students displayed modest levels of confidence in their abilities, but outperformed their more confident non-Asian peers.

For anyone who has ever spent a significant period of time in Japan or with Japanese people, I don't believe the above results will be all that surprising, considering the modest nature of many Japanese people compared to those from western cultures. Certainly there are exceptions, however, generally western culture tends to encourage individuals to view themselves positively, whereas Japanese culture much of the time, tends to encourage people to identify their weaknesses so that they can make efforts toward improving themselves. North Americans, for example, are generally raised to be accepting of information showing that they are better than others and cautious to believe information that presents them as worse than others. Japanese people, in contrast, are often not expected to exhibit this preference for self-enhancing information. In fact, given the importance of identifying and working toward eliminating negative information about the self, Japanese might be more inclined to accept self-relevant information that is negative.

Next, Mau (2000), in her study of decision-making self-efficacy and decision-making styles (rational, dependent, and intuitive) of Taiwanese and American college students, attributed the lower efficacy beliefs of the Asian students to "the collective-oriented culture (that) may have influenced Taiwanese students to rely less on individual abilities than on group efforts" (p. 374). Mau's study found that although a majority of American and Taiwanese students endorsed a rational style of decision-making (e.g., "I am very systematic when I go about making an important decision"), the Taiwanese chose a dependent style ("When I make a decision it is important to me what my friends

think about it”) as the second most likely choice, while the Americans were significantly less likely to endorse a dependent style. The author attributed the difference in decision-making style and in levels of self-efficacy to the Asian emphasis on social conformity and collective decision-making.

Similar to Eaton and Dembo (1997), Mau discusses the “modesty” of the response style of the Taiwanese students. While many European Americans display a tendency to “self-enhancement,” Asians display a tendency for “self-criticism,” which Mau claims is culturally adaptive because it confirms membership in the social unit. Although Mau acknowledges that Asian research respondents can display a response bias in which they typically provide lower self-ratings for most self-constructs in most settings, she states “The culture that is individual-oriented is more conducive to fostering self-efficacy, while the collective-oriented culture may have inhibited the development of self-efficacy” (2000, p. 374).

Finally, in a study comparing levels of general perceived self-efficacy among 13 cultural groups, Schwarzer and Born (1997) found that Japanese participants rated their general self-efficacy the lowest of the groups on the 10-item, 4-point scale, with a mean item rating of 2.02. Chinese participants were next lowest with an item mean rating of 2.31. In an extension of this research, Scholz et al. (2002) examined the same scale with a sample of over 19,000 participants from a total of 25 countries. Of the 25-countries and cultural groups, the presumably more collectivist participants from Japan (2.02) and Hong Kong (2.31) again rated general efficacy lower than other groups, while Costa Ricans (3.32), Danes (3.29), and French (3.22) participants rated generalized efficacy beliefs the highest. Although cultural dimensions were not assessed in this pair of studies,

the authors speculate that people from collectivist cultures might value hard work and effort more than perceived efficacy (Scholz et al., 2002).

Despite overwhelming evidence that self-efficacy judgments might differ slightly in collectivist societies, Bandura argues nonetheless, that self-efficacy whether it operates differently or not, absolutely still exists in such cultures. Self-efficacy theory, with its emphasis on personal control claims that “people guide their lives by their beliefs of personal efficacy” and “beliefs of personal efficacy constitute the key factor of human agency” (Bandura, 1997, p. 3), is firmly seated in a Western, independent, individualist context. However, Bandura (1997, 2002) rejects the notion that self-efficacy plays a lesser role in collectivist cultures: “People live their lives neither entirely autonomously nor entirely interdependently in any society. Interdependence does not obliterate a personal self” (1997, p. 32). He points out that groups within the collectivist dimension vary greatly, and that individuals, too, adjust their behavior depending on the context.

It is clear from the above research that efficacy beliefs operate differently in non-Western cultures than they do in Western cultures. In almost all of the studies that included direct comparisons of levels of efficacy beliefs, whether the studies compared pairs of cultural groups (e.g., Eaton & Dembo, 1997) or a large number of cultural groups (e.g., the 25 countries included in the study by Scholz et al., 2002), self-efficacy beliefs were typically higher for participants from Western, individualist cultures than for the participants from Asian or collectivist settings. The evidence also suggests that among collectivists, efficacy beliefs are typically lower but equally or even more predictive of performance, and that the calibration of their efficacy beliefs and subsequent functioning can be more accurate than among individualists.

With the above research in mind, faced with a self-efficacy questionnaire where they must rate their own perceived ability over a wide range of skills, it is no longer shocking at all that the Japanese teachers in this study might have tended to answer the questionnaire more modestly than their native English speaking counterparts, leading to a lower average profile as seen in Figure 12. Noteworthy in the above results, is that it is not that the Japanese believe that their efficacy levels are actually lower than they truly are, but rather that they are being modest in responding to the items on the questionnaire. It would be a huge mistake to assume that the Japanese population could not classify or distinguish their own levels of self-efficacy, however, the discrepancy in self-efficacy profiles compared to those of native English speakers can be explained by a more general modesty trend ingrained in a Japanese or collectivist society that was displayed in the profile analysis results.

In conclusion, the research cited above leads us to several plausible reasons why in this study, such a gap exists between the native and non-native teachers self-efficacy beliefs, including the modest demeanor of many Japanese citizens, the difference in individualistic versus collective cultures, and the self-perceptions of non-native English teachers. However, the results simultaneously lead us to the interesting question of whether Japanese English teachers' sense of efficacy, their beliefs in their capability to successfully handle specific teaching situations, is at all affected by this dichotomy of nativeness or influenced by their cultures. As we know, teacher self-efficacy is not about who is actually a good or better teacher but rather about who believes that they are good. Whether this can be accurately assessed by a singular questionnaire remains to be seen. Nonetheless, this is extremely important and certainly worthy of future research because

as Bandura (1997) has addressed, teachers' self-efficacy can ultimately impact the environment teachers create in their classroom, their specific instructional practices, their orientation toward the educational process, and, perhaps most importantly, their students' academic achievement.

Research Question 2

The second research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' years of experience. This question was answered by conducting a profile analysis, where the teacher's years of experience (i.e., 0-10, 11-20, and 21+) were entered as the independent variables and the Rasch person ability estimates of the four self-efficacy factors obtained from the questionnaires were entered as the dependent variables.

The answer to the primary question, "do the teachers with different years of experience in the study have parallel profiles," was yes, indicating a non-significant effect. In other words, this test revealed that the three groups of teachers had no significantly different patterns of highs and lows when the means of the self-efficacy scores were compared. The same up-down trend of the lines for both native and non-native English speakers can be seen in Figure 13, and the fact that the lines have very similar trajectories indicates a non-significant effect. The answer to the second question, "are the three groups at equal levels across observations", was no, indicating a significant effect. Essentially this test showed that one or more groups scored higher on average across all measures. Upon closer examination, the results indicated that teachers with 11-20 years of experience as well as 21 or more years of experience, on average, have

reliably higher self-efficacy scores than teachers with 0-10 years of experience. That is to say, more experienced teachers rated themselves significantly higher than the less experienced teachers on all four self-efficacy factors. This test is displayed in Figure 13 by the fact that the 11-20 year profile and the 21 years or more profile are separated by a considerable gap from the 0-10 year profile. The answer to the third question, “do the profiles exhibit flatness across observations”, was also no, indicating a significant effect. Because the profiles did not exhibit flatness across observations, it suggests that teachers’ self-efficacy varied regardless of whether the teacher was experienced or not. In Figure 13, this is shown by no two points on the same line to have the same score.

The first thing that stands out when looking at Figure 13 is that the measured self-efficacy variables are noticeably higher for the two more experienced group of teachers (11-20 years and 21+ years) than for those teachers in the first group (0-10 years). This suggests that on average, the more experienced teachers perceived themselves to have greater ability to engage their students, superior instructional strategies, better classroom management techniques, and better ability to deal with their superiors. From a self-efficacy standpoint, these results are not surprising and are completely inline with the current research. Bandura (1977, 1997) believes mastery experience is the most influential source of self-efficacy, implying that direct hands-on success and accomplishment can help develop a strong sense of efficacy. As is inferred in Bandura’s triangular model of reciprocal determinism (see Figure 1 in Chapter 1), individuals engage in tasks and activities, interpret the results of their actions, use the interpretations to develop beliefs about their capability to engage in subsequent tasks or activities, and act in concert with the beliefs created. Typically, outcomes interpreted as successful raise

self-efficacy; those interpreted as failures lower it. And because throughout a university English teacher's long career, they are exposed to a wide variety of experiences and have at least some successful encounters, the assumption is that more experienced teachers have a higher likelihood of believing themselves to be more efficacious (Karimvand, 2011).

Additionally, Bandura (1997) hypothesized that although self-efficacy beliefs remain relatively stable once established, researchers have noted, little evidence exists about how efficacy beliefs change or solidify across stages of a career. Moreover, some studies have shown a positive correlation between years of experience and efficacy beliefs of teachers, in which seasoned teachers reported higher levels of self-efficacy than their novice counterparts (Lin & Tsai, 1999; Liu et al., 2007; Wolters & Daugherty, 2007). For example, in the recent study by Wolters and Daugherty (2007), they used a large online sample of teachers (N =1,024) from the United States to examine the influence of teaching experience on teachers' self- efficacy. Teachers were divided into four experience groups: 1 year, 1-5 years, 6-10 years, and 11+ years of experience and completed the Teachers' Self-Efficacy Scale by Tschannen-Moran and Woolfolk Hoy (2001). Results showed modest effects of experience on self-efficacy for instructional strategies ($\eta^2 = .04$) and self-efficacy for classroom management ($\eta^2 = .02$), but no effect of experience on self-efficacy for student engagement. Although Wolters and Daugherty's research has made important contributions by linking experience with teachers' self-efficacy, the findings paint an incomplete picture, with two potential limitations that are relevant to the current study. One problem is that the relationship between teachers' self-efficacy and experience might not be linear. In other words, they

found that teachers' self-efficacy initially rose and then fell over three data collection points at the beginning of teachers' careers. As the current study is not of a longitudinal nature and only reveals a cross-section of the university English teachers' career, unfortunately it cannot speak to a possible non-linear relationship between self-efficacy and experience.

Another problem with the above study is that teachers with more than 10 years of experience were treated as a single group. The authors of the study acknowledged that the "lack of differentiation among the most-experienced teachers may mask changes in teachers' self-efficacy that may occur toward the end of their careers" (p. 189). In the current study however, the teachers were split into three groups and in fact the two most experienced groups (11-20 years and 21+ years) showed very similar profiles, perhaps lending support to the idea that changes to the most experienced teachers' self-efficacy is not being "masked."

When looking a bit more closely at the results, although the experienced teachers in this study were found to have higher self-efficacy beliefs than the less experienced teachers on all four subscales of the JULTEBS, it is interesting how low the scores were for all groups of teachers in the *Efficacy for Dealing with Superiors* subscale. One explanation might be that no matter the level or years of experience, dealing with one's superiors, the various personalities that exist, and the ins and outs of the administration, is a task that while perhaps becomes familiar over time, is not a realm where one can build their confidence easily.

Additionally, the low scores found in the *Efficacy for Dealing with Superiors* subscale were also mirrored quite consistently by the comments of teachers from all

levels of experience whom I interviewed. For example, both Yoshi, a tenured 33 year old teacher with only 6 years of experience, professed:

It's really difficult to get what I believe to be simple things done quickly here. There is always a long drawn out process involved. Members of the administration are not always on the same page as us and I'm not always sure who has spoken to who. It's a bit confusing.

Yoshi appeared to be equally as frustrated with the difficulties in dealing with his superiors as Kenta, a tenured 57-year-old teacher with 25 years of experience, who explained to me:

Nobody communicates either. There's not enough talk between teachers and the administration, and therefore nobody knows what's going on. You always have to fight to get something changed. It's really exhausting!

Certainly there are bound to be exceptions, but what is plainly evident is that regardless the years of university teaching experience, dealing with one's superiors is far from a task many teachers hold much confidence. Furthermore, it also seems to be an undertaking that might be impervious to developing any amount of self-efficacy.

In conclusion, the results of the second research question make sense logically. As mentioned above, of the four major influences on teachers' self-efficacy beliefs, the most powerful is mastery experiences, which for university English teachers in Japan comes from actual accomplishments inside as well as outside of the classroom. Efficacy beliefs are raised if an English teacher perceives his or her teaching implementation to be a success, which then contributes to the expectations that future routines will likely be positive as well. Therefore, it might be expected that teachers, who experience this success more often over the years, will naturally show greater levels of self-efficacy, just as the experienced teachers did in this study.

Research Question 3

The third research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' tenured or limited-contract status. This question was answered by conducting a profile analysis, where the teacher's contract status was entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors obtained from the questionnaires were entered as the dependent variables.

The answer to the primary question, "do tenured English teachers and limited-term contract English teachers in the study have parallel profiles," was no, indicating a significant effect. In other words, this test revealed that the tenured teachers and limited-term contract teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared. Although a similar up-down trend of the lines for both native and non-native English speakers can be seen in Figure 14, the fact that the lines have marginally different slopes, especially from points two to three, and from three to four, suggest a significant effect. The answer to the second question, "are the two groups at equal levels across observations", was yes, indicating a non-significant effect. These results indicated that tenured teachers, on average, do not have reliably higher or lower self-efficacy scores than limited-term contract teachers on the selected set of measures. That is to say, both groups of teachers rated themselves similarly on all four self-efficacy measures. The results of this test can be seen in Figure 14 by the fact that the lines are right on top of each other, no statistically significant gap exists between points, and the profiles even cross over between points three and four. The answer to the third question, "do the profiles exhibit flatness across observations,"

was no, indicating a significant effect. This suggests that teachers' self-efficacy varied regardless of whether the teacher was tenured or on a limited-term contract. In Figure 14, no two points on the same line having the same score visually demonstrate this result.

There are several interesting points found in the results of the third research question. To begin with, it is obvious in Figure 14 that the tenured teachers as well as the limited-term contract teachers rated themselves relatively equally across the four self-efficacy factors, with the contract teachers rating themselves slightly higher than the tenured teachers on all factors except the *Dealing with Superiors* factor, where we see the largest discrepancy. It should come as no surprise that the levels of self-efficacy that both the tenured and non-tenured groups of teachers feel toward *Student Engagement*, *Instructional Strategies*, and *Classroom Management* are nearly the same. This is because much like the race or gender demographic, holding tenure or not holding tenure alone will theoretically have little bearing on how one perceives their own abilities and how confident they are in attaining desired outcomes in the actual classroom. Other factors, such as student level, teacher training, and years of teaching experience (Bandura, 1997), will influence self-efficacy in the classroom to a much greater degree.

On the other hand, in the *Dealing with Superiors* factor, the tenured teachers rated themselves considerably (i.e., nearly one logit) higher than the contract teachers. In this case, holding tenure or having a limited-term contract seems to strongly affect how confident one is when dealing with superiors and administration outside the classroom. Although the differences found are neither statistically significant nor, as can be seen by their low scores, have either of the two groups of teachers judged themselves to be

particularly efficacious in *Dealing with Superiors*, the ensuing discussion will focus on the disparity between how the two groups rated themselves.

In order to illustrate why the tenured teachers as a whole, displayed higher average self-efficacy scores for *Dealing with Superiors* than non-tenured teachers, I'll highlight four possible challenges and impediments that exist for a typical limited-term contract English teacher in Japan. The following challenges are similar to those discussed in Chapter 6 and are adapted from Murray's (2013) article on teacher burnout in higher education:

1. In addition to feeling the stress of guaranteed unemployment after a set number of years, many contract teachers might begin to feel detached, particularly from other faculty members in their departments. Although a contract teacher works full time, they are often fully aware of both the indirect and blatant distinctions that are continually made between them and tenured faculty. A simple example could be a departmental meeting that is only for tenured faculty members. Many contract workers are not involved in any decisions that have a direct impact on them and their students despite having made a substantial commitment in terms of time and energy to their universities.
2. Another impediment for many contract teachers is salary. Although most limited-term instructors receive a decent salary, many do not receive annual bonuses nor do they qualify for severance packages upon retirement. Additionally, many institutions do not provide a research budget for professional development expenses such as academic conferences and professional associations for contract teachers, but do for their tenured counterparts.
3. Many contract teachers have a comparatively heavy course load. They are often

required to teach ten, even twelve, 90-minute courses each week. Furthermore, it is often the case that these teachers have little input into the course offerings. They often are stuck with larger compulsory classes and do not teach seminar courses like their tenured colleagues. Perhaps the most aggravating point is that often all major aspects of the courses they teach (textbook, syllabus, and evaluation methods) have been set by the administration.

4. Another challenge for contract teachers is that often their jobs do not allow them to focus their efforts on personally meaningful activities. This means that although tenured faculty are expected to make contributions to the school in a number of areas, contract teachers' primary responsibility is to teach compulsory freshmen and sophomore classes. Despite their commitment to scholarship, often institutions do not give him formal opportunities to be involved in research activities.

In addition to the four examples above, as one interview participant declared, "guaranteed unemployment" is yet another heavy burden that the limited-term contract English teachers must bear. The burden can be so great in fact that it is completely reasonable to deduce that in some cases, the emotional stress that a limited-term contract situation triggers could create such a toxic environment, that actually not having a tenured position, could easily lower self-efficacy beliefs when it comes to dealings with superiors. This is of course what Bandura (1997) meant when he explained that a physiological state, or one's mood, affects people's judgments of their personal efficacy. He stated that just as a positive mood augments self-efficacy, a hopeless and unhappy mood weakens it. In other words, self-efficacy beliefs will deteriorate faster if people's stress is increased and their negative emotional tendencies are fostered. Working as a

teacher with an uncertain future and a definitive end of employment quickly approaching, could easily generate a mood or atmosphere that is potentially incapacitating as well as preventing the growth of self-efficacy beliefs.

Furthermore, there is also a psychological component that might help explain why tenured faculty rated themselves more self-efficacious when dealing with superiors. With tenure comes the previously discussed job security as well as academic freedom. However, obtaining tenure can also deliver another benefit: the belief of belonging to a group. Everybody has a superiors, even tenured faculty have superiors. However, there is undeniably a greater feeling of equality among the staff members who hold tenure compared to those who are on limited-term contracts. With a secure job, there is less to lose, there is less holding them back, there is less worry about expressing one's opinion over fear of what others will say or think about them. A limited-term faculty member does not have the luxuries just mentioned. A limited-term contract teacher could be a more effective teacher, better liked by his students, and respected by his peers than a tenured teacher, but undoubtedly, the majority of the time the tenured faculty, despite any narrow-mindedness that a superior might display, feel less timid in their interactions with a superior. In other words, there is a sense of comfort in a tenured position, where the limited-term contract teachers instead feel instability. Finally, it is perfectly reasonable to believe that many limited-term faculty members, especially the younger teachers, could hold a sense of inferiority towards their superiors, especially if these superiors are tenured.

Consequently, the results from the third research question displayed above that show the tenured faculty having a greater self-efficacy in *Dealing with Superiors* than

their limited-term contract teacher counterparts are not extraordinary by any means, but point to a real issue that is affecting real English teachers at Japanese universities.

To conclude the discussion about Research Question 3, in a post-interview follow-up e-mail exchange with Robert, a contract associate professor at a national university, he made clear to me yet another curse of contract teaching positions at Japanese universities in light of new legislation passed (Rivers, 2013) limiting teachers to five-year contracts. I think it sums up quite well the myriad of difficulties that impermanent faculty face.

Mandatory retirement, especially at the ages commonly designated by many universities, is an odd and particularly unfair practice to maintain, particularly considering Japan's age demographics. It is unquestionably tough for many perfectly able-bodied workers with plenty of productive years left in them. However, which would you rather be, a 50-something limited-term contract worker who has worked securely on year-to-year contracts for a long time suddenly learning he will be fired within 5 years in a country where age discrimination is legal and widely practiced, therefore genuinely threatening him with the prospect of unemployment and no solid retirement to aid him at age 55? Or a 50-something full-time contract worker who has previously been given verbal assurances that extensions to his contract are pretty much given, but then learning that he is abruptly placed on a 5-years-and-you're-out contract? Add the fact that in both of these cases the teacher is also a substantial or sole breadwinner in a family with children (as is the case with many people I know who might be affected by these changes). Add huge expenses of time and money invested in higher education to improve knowledge and job security, now mostly irrelevant, because everyone, regardless of backgrounds, skill, publication record, motivations, Japanese ability, etc., is automatically fired in five years. Would you prefer to have one of these positions or to be a 50-something with tenure, not facing forced retirement until 65 or even later in some circumstances? I know which bucket most people would prefer.

Research Question 4

The fourth research question asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' gender. This question was answered by conducting a profile analysis, where the teacher's gender (i.e., male or female) was entered as the independent variable and the Rasch person ability estimates of the four self-efficacy factors obtained from the questionnaires were entered as the dependent variables.

The answer to the primary question, "do male and female English teachers in the study have parallel profiles," was yes, indicating a non-significant effect. In other words, this test revealed that male and female teachers had the same patterns of highs and lows when the means of the self-efficacy scores were compared. A very similar up-down trend of the lines for both groups can be seen in Figure 15. The answer to the second question, "are the two groups at equal levels across observations", was also yes, indicating a non-significant effect. This result indicated that male teachers, on average, do not have reliably higher or lower self-efficacy scores than female teachers on the selected set of measures. Figure 15 displays these results with the two lines right on top of each other, with no statistically significant gaps existing between points. The answer to the third question, "do the profiles exhibit flatness across observations," was no, indicating a significant effect. These results suggest that teachers' self-efficacy varied regardless of whether the teacher was male or female. In Figure 15, this is shown by no two points on the same line to have the same score.

Although demographic variables such as gender and race have typically not been strong predictors of the efficacy beliefs of teachers (Bandura 1997), with Japanese

university faculty and staff member demographics being what they are, predominantly male (Japanese Association of National Universities, 2013), gender was hypothesized to be a variable that could influence a teacher's professional life. However, in line with the current self-efficacy research and as is displayed in the results of the third research question, the demographic variable, gender, was found to not be systematically related at all to the self-efficacy beliefs of the English teachers in this study.

Furthermore, Bandura's social cognitive theory does not endow gender with motivating properties (Bussey & Bandura, 1999) and theorists contend that environmental factors and personal factors other than gender are at work in creating motivational and self-regulatory differences in individuals. This would appear to be the case in this study as well, as the differences in scores on the JULTEBS between the men and women are virtually indistinguishable.

Additionally, researchers have also observed that academic fields such as mathematics, science, and technology, are usually classified within a male-domain. In these areas, a masculine orientation is associated with confidence and achievement because masculine self-perceptions are filled with the notion that success in these areas is a masculine imperative (Eisenberg et al., 1996). Conversely, language and literacy are typically viewed as being female-domains (Eccles, 1987). As a result, a feminine orientation is associated with motivational beliefs related to success in skills related to languages. Although not the case in this study, the above research would imply that if one gender group were to have a so-called "advantage" and likely have greater confidence, it would be the women. Obviously, the results showed no such advantage for either the men or the women.

In addition to the research cited above, Zeldin and Pajares (2000) explored the personal stories of women who excelled at careers in areas traditionally dominated by males to better understand the ways in which their self-efficacy beliefs influenced their academic and career choices. They found that the messages the women received from significant others in their lives, as well as the vicarious experiences they underwent, nourished their self-efficacy beliefs as they set out to meet the challenges required to succeed in male-dominated academic domains. Their findings suggested that girls developed higher self-efficacy beliefs and engaged more self-regulatory strategies in homes and classrooms in which parents and teachers stressed the importance and value of academic skills, encouraged girls to persist and persevere in the face of academic and social obstacles, and broke down stereotypical conceptions regarding academic domains. This particular angle was not explored in this study, but would be especially interesting to explore today, because there has been a very public and open push at many universities to actively promote the hiring of more women faculty members (Japanese Association of National Universities, 2013).

Despite Bandura stating that gender does not have a strong influence on teachers' self-efficacy, upon reviewing studies that have focused on the relationship between gender and teachers' sense of efficacy, there are in fact a number of discrepancies that can be seen. Imants and De Brabander (1996), using a modified version of the Teacher Efficacy Scale (TES), concluded that gender does in fact influence teachers' self-efficacy. More precisely, their results showed that male elementary school teachers' level of efficacy beliefs for pupil-oriented and school-oriented tasks, seemed to be higher than their female counterparts. In complete contrast, both the work of Cheung (2008) and

Karimvand (2011) in fact found that female teachers have significantly more general efficacy than male teachers. However, it is necessary to note that many of the female teachers in the studies by both Cheung and Karimvand were either older or had more teaching experience than their male teaching counterparts, actually lending support to the idea that more experienced teachers have higher self-efficacy as discussed earlier in this chapter. Finally, and in line with the results of the current study, there are yet even other studies, such as Ghaith and Shaaban (1999) and Tschannen-Moran and Woolfolk Hoy (2002), which showed gender had absolutely no significant effect on teachers' self-efficacy.

In conclusion, despite several studies both leaning in favor and against gender influencing teachers' self-efficacy beliefs, there are two conclusions that can safely be made. The first is that the current study showed no statistically significant differences between male and female English faculty members' self-efficacy over the four subscales on the JULTEBS. These results are inline with Bandura's assessment of the gender demographic not having an influence on teachers' self-efficacy. The second fact is that there is no consistent trend or general direction in the current research where gender might correlate with teachers' efficacy beliefs, as too many discrepancies and opposing observations still exist. This being the case, further investigation is required to definitively determine gender's influence on teachers' self-efficacy.

Contributions and Implications

The present study provides both contributions to the field of teacher self-efficacy as well as implications for teachers and policy-makers in the Japanese university English education context. Below, I discuss these points based on the findings of this study.

The current study was significant in several ways and made various contributions to expanding the current state of knowledge in the TESOL field. First, the study addressed the need to inquire into university teachers' beliefs in their capabilities for teaching English by adopting the notion of teacher self-efficacy. Although many studies had documented the formidable influence of teachers' sense of efficacy in teaching and learning, very few researchers had studied it in the TESOL field. Moreover, no researchers had investigated university teachers' sense of efficacy in an EFL setting in Japan. This study proved valuable by investigating the levels of teachers' self-efficacy to provide a current depiction of university English education in Japan from the teachers' perspective.

Furthermore, because self-efficacy is domain specific, the exploration of self-efficacy beliefs in the Japanese university context made a valuable contribution to the broader field of self-efficacy research. This study introduced a new theoretically and psychometrically sound instrument, the Japanese University Language Teachers' Efficacy Beliefs Scale (JULTEBS), to the list of available self-efficacy inventories, and because of the thorough validation methods used (i.e., the Rasch rating scale model and a confirmatory factor analysis), will be appealing to those researchers involved in self-efficacy as well as general scale development.

In addition to the previous points, this study is also significant because of the

methodology employed. There had been a surfacing need for qualitative examination in addition to quantitative examination in teacher efficacy research. The design of the current study (a triangulation strategy mixed-method design) addressed this point, as can be seen with the interviews helping to create a richer depiction of what potentially supports and challenges teacher self-efficacy beliefs in Japan.

Perhaps the greatest contribution this study has to offer, deals with the difficulties and challenges that were identified by the teachers in the semi-structured interviews and on the questionnaires. The difficulties centered on dealings with superiors and administrative policies. Because every institution of higher education in Japan today employs numerous English teachers and offers a variety of English course options. The JULTEBS, a standardized scale that evaluates the self-efficacy beliefs of English instructors, if administered in a non-threatening way, could provide the administrators and policy makers the means to identify where improvement and support is needed. If the proper infrastructure is established, professional development programs, faculty mentoring, instructional workshops, and more ideal working conditions could be developed to improve English teaching at Japanese universities. In addition, the administrators could gain a greater understanding as to what components need to be included in graduate education and training programs. This understanding would provide a means through which postsecondary institutions and graduate programs could better prepare future faculty for university English teaching roles. Ultimately, the JULTEBS will diagnose potential challenges and difficulties so efforts toward solutions and better educational practices can be pursued.

Next, according to Social Cognitive Theory (Bandura, 1997), one's belief about

their own ability to perform a certain task is critical because how people perform can often be better predicted by the beliefs they hold about their capabilities than by what they are actually capable of accomplishing. Therefore, given the importance of teacher's sense of efficacy, it seems critical to not only prepare teachers with strong positive beliefs about their capabilities early on, but also create an environment in the university where self-efficacy can prosper. Teacher development programs and better communication between colleagues, superiors, and the administration should be augmented in order to seek ways to promote teachers' efficacy levels.

Finally, although there is no single strategy or magic button that will immediately raise the self-efficacy of language teachers at Japanese universities, the following are three practical motivational strategies and suggestions that English languages teachers can do on their own to help boost their teaching efficacy.

A teacher's successful emotional management can improve moods and thinking not only for themselves, but also for those they interact with. In the workplace, according to one study (Diefendorff, Richard, & Yang, 2008), here are the top ten most commonly used emotion regulation strategies, ranked by frequency of use: (1) Seek out individuals who make you feel good, (2) keep yourself busy working on other things, (3) do something enjoyable to improve your mood, (4) try to solve the problem, (5) find humor in the situation, (6) think about how the other person feels, (7) consider how things could be worse, (8) pretend you are in a good mood, (9) turn your attention to something that doesn't bother you, and (10) remind yourself that you cannot control everything.

Asking teachers how they successfully cope. Their stories and encouragement (i.e., vicarious learning and social persuasion) might help create positive emotions and

motivate other teachers to keep learning in new ways. Another critical emotion for enabling well-being in difficult times is hope. Hope generates agentic and pathway thinking, the cognitive processes of believing in the ability to maintain self-control toward achieving one's goals, and imagining alternative routes toward them when the way is blocked (Snyder, Cheavens, & Sympson, 1997). Without hope, individuals neither see nor seek ways around obstacles. And with it they become capable of overcoming obstacles and engaging toward their preferred futures (Carver & Scheirier, 2007).

As discussed in Chapter 2, depersonalization was found to be the second major contributor to teacher burnout (Grayson & Alvarez, 2007). The sense of social support is crucial for teachers under stressful conditions (Kieschke & Schaarschmidt, 2008). This is in-line with Bandura's (1997) theory that social persuasion is one of the four main influences to self-efficacy. With joining a community, teachers can find the relatedness that they need to persist and even to maintain well-being. Relevant here is the belonging hypothesis, which is that "human beings have a pervasive drive to form and maintain at least a minimum quantity of lasting, positive, and significant interpersonal relations" (Baumeister & Leary, 1995, p. 497). People are able to become motivated through socialization.

Communities can enrich one's life through the interpersonal processes of forming and attaining both personal and mutual goals (Snyder, 1994). Through observing and interacting with others, one learns ways of doing things that they wouldn't have figured out alone. This is the same as Bandura's (1997) theory that vicarious learning is the second most influential source of self-efficacy. One can increase the ability to cultivate pathways for attaining individual goals alongside interrelated goals when enacted through

the synergy of a cooperative community. For example, teachers can collaborate to mentor each other, and model the behaviors and values that lead to professional success (Hooker, Nakamura, & Csikszentmihalyi, 2003). Meeting to talk about classroom problems and solutions, forming reading circles that focus on pedagogy and research, and collaborating on projects are all possible ways to promote higher self-efficacy in teaching.

As explained throughout this study, teacher self-efficacy is the underlying psychological force that drives teachers toward self-directed professional development that can lead to job satisfaction (Mushayikwa & Lubben, 2009). Another suggestion to boost professional self-efficacy is committing to a long-term goal and dividing it into attainable sub goals. With each achievement, one can build self-efficacy, impelling persistence, improving outlook and demeanor, and increasing chances of getting the recognition needed to help take control of one's career development. The key is maintaining hope through agentive and pathway thinking.

An example of a sub goal might lead from creating materials for class, to an article in a journal, to a textbook series. One's in-class methods development could step from an academic presentation to curriculum implementation at a school (Falout, 2010). A language teacher can begin this process right from where they are currently employed. They can ask students what went well with class. Ask teachers and administrators about the needs of the institution. Keeping up with the evolving needs of students, institutions, and society boosts self-efficacy. Moreover, when teachers see progress in themselves, others can too, perhaps even recognizing those teachers as valuable members of the community.

Teachers can practice self-regulatory strategies to maintain the motivation to persist and improve their teaching situations and careers. Wishful thinking won't work and pathways can only form when teachers vividly imagine and act upon them. Managing emotions, joining communities, and boosting efficacy are three ways to promote proactive step taking toward preferred futures for a teacher's students, colleagues, and themselves.

In the next chapter, I present a final summary of the study, limitations of the study, future research that should be conducted in this area, as well as my overall conclusions.

CHAPTER 8

CONCLUSION

Summary of the Study

There were two main objectives that were achieved in this study. The first was to introduce and validate an instrument that did not previously exist, appropriate for investigating Japanese university English teachers' self-efficacy beliefs. The second was to use the instrument to explore the self-efficacy beliefs of university English teachers in Japan, a subdivision of the population that had never been previously studied. The following is a summary of my results.

Before conducting the main analyses, I examined the questionnaire data using the Rasch rating-scale model as well as a confirmatory factor analyses in order to verify the validity and reliability of the questionnaire. After a comparison and careful consideration of the results of both analyses, misfitting and outlying items as well as participants were cut from the study. In the end, 392 participants completed the Japanese University Language Teacher Efficacy Beliefs Scale (JULTEBS), a 28-item self-efficacy questionnaire. Next, in order to answer the first four research questions, I performed four profile analyses, a special application of a MANOVA, to compare the various groups of participants by their average scores across the four self-efficacy factors, Student Engagement, Instructional Strategies, Classroom Management, and Dealing with Superiors. Finally, in order to answer the last two research questions, I conducted semi-structured interviews with 12 participants in the study to gain a more complete picture of English teacher self-efficacy at Japanese universities.

Research question 1 asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' native language. Results from the profile analysis revealed that native English teachers and Japanese English teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared and that native English teachers, on average, have higher self-efficacy scores than native Japanese teachers on the selected set of measures. Finally, the results suggested that teachers' self-efficacy varied regardless of whether the teacher was a native English speaker or native Japanese speaker.

Research question 2 asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' experience. Results from the profile analysis revealed that the three groups of teachers had no significantly different patterns of highs and lows when the means of the self-efficacy scores were compared and that more experienced teachers, on average, have reliably higher self-efficacy scores than less experienced teachers on the selected set of measures. Finally, the results suggested that teachers' self-efficacy varied regardless of whether the teacher was more or less experienced.

Research question 3 asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' tenured or limited-term contract status. Results from the profile analysis revealed that the tenured teachers and limited-term contract teachers had significantly different patterns of highs and lows when the means of the self-efficacy scores were compared but that tenured teachers, on average, did not have reliably higher or lower self-efficacy scores than limited-term contract teachers, on the selected set of measures. Finally, the results suggested that teachers' self-efficacy varied

regardless of whether the teacher was tenured or on a limited-term contract.

Research question 4 asked to what extent Japanese university EFL teachers' self-efficacy beliefs differ according to the teachers' gender. Results from the profile analysis revealed that male and female teachers had the same patterns of highs and lows when the means of the self-efficacy scores were compared and that male teachers, on average, did not have reliably higher or lower self-efficacy scores than female teachers, on the selected set of measures. Finally, the results suggested that teachers' self-efficacy varied regardless of whether the teacher was male or female.

Research questions 5 and 6 asked what could strengthen and weaken university English teachers' self-efficacy. The results of the semi-structured interviews yielded a total of seven emergent themes. The first four themes, (Autonomy, Colleagues, Money, and Students) addressed qualities that could support teachers' self-efficacy. The final three themes (Administration, Students, and Limited-term Contracts) were qualities that could weaken teachers' self-efficacy.

While it is my hope that these findings provide awareness in to inquiries regarding self-efficacy and teacher-self-efficacy in Japan, this study has limitations that I discuss in the next section.

Limitations of the Study

The first limitation concerns the items from the questionnaire, particularly those that measured efficacy in Student Engagement, Instructional Strategies, and Classroom Management. Some of the items from these three self-efficacy factors, despite fitting the Rasch model and achieving all the recommended benchmarks proposed by Linacre

(2009), demonstrated a ceiling effect. A comparison of the locations of the means for the person ability measures and item difficulty measures showed that the mean person measures were higher than the mean item measures. This indicated that the participants easily endorsed the items. In other words, they believed that they were highly capable of carrying out the tasks described by the items. Ideally, the mean person ability and mean item difficulty should be as close as possible, which would demonstrate that the items on the whole were well targeted for the sample. In light of the ceiling effect obtained in the current study, consideration should be given to making some of the items measuring each factor more difficult in order to achieve a more balanced instrument.

The second limitation concerns the concept and word choice of “superior” versus “administration” used in the items measuring Efficacy in Dealing with Superiors. There was a suspected problem with how different teachers in different teaching situations would interpret the use of one word over the other. During the initial development and pilot of the questionnaire, both the terms superior and administration were used somewhat interchangeably. However, in order to try to eliminate confusion and create a more unidimensional construct, the term superior was chosen over administration, and all items were adjusted accordingly to represent the change. However, through direct feedback from several participants in the study, it became clear that in certain teaching contexts, particularly those of limited-term contract teachers, direct superiors were not always known by the teachers and in some contexts, no actual communication ever takes place between the teacher and their superior. Instead, curriculum guidelines, textbooks choices, class schedules, research activities, etc... all appeared to be dictated and conveyed to the teachers by a general administration office. In such cases, it is possible

that the English teacher is in fact unaware of who is actually their direct superior. Consequently, it is likely that these participants had difficulty answering the items on the questionnaire designed to measure Efficacy in Dealing with Superiors.

Third, based on the results of the qualitative data and the various topics that arose in the interviews, there were variables that were not included in this study that might have been helpful in determining the levels of teacher self-efficacy. Furthermore, several of the following variables perhaps even warranted an entire additional factor on the JULTEBS, used in this study. These factors include requirements for conducting research, school level and rank, whether or not the teachers taught English majors or taught in an English department, the number of English teachers in the department, the average class level and size, whether or not the classes they taught were compulsory or self-designed seminars, the quality of the facilities, and the length of time they have been working at their current university. Any of these factors could help paint a more complete picture of how teachers perceive their abilities and their inclusion in future studies warrant consideration.

Although every attempt to constrain the deleterious effects that these limitations might have posed to the results of the study, the findings derived from the various analyses should be interpreted cautiously. Despite the above limitations, based on the summary of the study and the current relevant literature, I would like to offer suggestions for future research.

Suggestions for Future Research

The present exploratory study adopted a mixed methods design to investigate Japanese university English teachers' self-efficacy beliefs. The study revealed how

different groups of teachers perceived their own abilities to complete desired tasks. It further examined what factors potentially help as well as hinder teachers' confidence. I hope that the key findings of this study will contribute to the fields of self-efficacy and EFL education. However, there is additional research in this area that can shed light on important issues. It is my hope that future researchers pursue the following topics.

First, more research using the JULTEBS developed for the present study is needed for comparison purposes. As the first effort adopting a subject-specific approach in measuring teachers' sense of efficacy in the Japanese university EFL education context, the present study has the limitation that the results cannot be generalized beyond its research participants. Accordingly, more studies with different samples would provide useful information in understanding the current levels of Japanese university English teachers' self-efficacy beliefs.

Second, more research is needed to investigate how one's English teaching efficacy is formed. The present study found themes that potentially could support or hinder the development of one's self-efficacy. Yet, it remains unanswered what process exactly contributes to one's confidence in teaching English at a Japanese university. Future research should focus on the four sources of efficacy beliefs, (i.e., mastery experiences, vicarious experience, verbal persuasion, and physiological states). For example, more comprehensive research could be conducted to examine how attitudes about one's English teaching colleagues or students influences their confidence in teaching English through observation and/or multiple in-depth interviews.

Third, more empirical assessment is needed of the sources of efficacy building information, collective teacher efficacy, and methods for impacting efficacy change in

English language teachers. Additionally, utilization of stronger measurement methodologies, including appropriate factor retention rules, confirmatory approaches, and higher order analyses, would likely enable more robust and parsimonious scales and inventories that generate more reliable and valid scores. It is also relevant to extend the traditional questionnaire form for measurement to include other methods of efficacy evaluation. For example, because an efficacy judgment is a result of an individual's filtering of internal and external factors, the context surrounding a person's judgment is very relevant to the study of teacher efficacy, but this is captured poorly by Likert-type questionnaires. Future research should consider the university context more directly via observation or perhaps attempt to evaluate teacher cognition by conducting a "think aloud," in which teachers can elaborate on why they responded the way they did and generate examples.

Next, a major part of a university teacher's overall workload in Japan, as well as many other countries, is their duty to conduct research, publish, and present their findings. This element was not explicitly explored in this study. It would be interesting and perhaps informative to include another sub construct to the current scale that measured 'efficacy in conducting research', in order to achieve a more complete picture of the self-efficacy beliefs of university English teachers.

Finally, perhaps the ultimate future research goal of teacher self-efficacy is to show that high teacher self-efficacy leads to superior student results. If higher teacher self-efficacy translated in to higher student proficiency scores, the importance of having a teacher with a strong sense of self-efficacy would be difficult to deny. This is of course not easily researched. In the context of the current study, ideally, the researcher would

need to follow a group of English teachers (at least 30 teachers from various institutions) and their students over the course of at least one year. The teachers' efficacy scores would be compared to the students' language proficiency scores at regular intervals throughout the school year. Essentially, statistically greater increases in student language proficiency taught by more efficacious teachers would provide great support for the reputation of self-efficacy. Such research lends itself to the use of Hierarchical Linear Modeling (HLM) because data for the participants in such a study would be organized at more than one level (i.e., nested data). In this case, the school level, the teacher level, and the student level. The units of analysis in HLM are usually individuals (at a lower level) who are nested within contextual units (at a higher level). While the lowest level of data in HLM is usually an individual, repeated measurements of individuals might also be examined (Tabachnick & Fidell, 2007).

Final Conclusions

The theory of teacher efficacy has been around for more than 35 years, but measurement of teachers' self-efficacy beliefs has still not come to fruition. As I mentioned in the introduction, educational researchers have attempted many times to measure teacher self-efficacy, however these attempts have been plagued with theoretical and measurement issues. Despite the growing body of evidence that past measures and scales are inadequate, confusion remains, and lackluster self-efficacy inventories and flawed methodology are a common appearance, even in well-respected journals. There is however evidence in the literature that progress is being made in creating theoretically sound measures of teachers' self-efficacy beliefs, but there is still definitely room for

improvement. Given these recent theoretical advances and better-founded measurement attempts, and assuming continued efforts are made to connect theory and measurement integrity, the study of teacher self-efficacy seems ready to thrive and prosper as a well-respected construct in educational psychology.

This study has tried to help teacher self-efficacy take a step forward by putting forth a new theoretically sound, reliable, and valid instrument in to the EFL and self-efficacy fields. I believe that this new measure, the JULTEBS, is not only acceptable but also worthy of future exploration. Additionally, the results obtained in the current study provide a firm starting point, in which further investigation in to the beliefs of English teachers can begin. Greater understanding of the self-efficacy beliefs of university language teachers has the potential to expand scholars' understanding of this construct as well as support teacher educators, administrators, and policy-makers in fostering teachers' self-efficacy. Finally, and perhaps most importantly, I believe that increased awareness of as well as a commitment to support teacher self-efficacy, has the potential to have an enormously positive effect on student outcomes.

REFERENCES

- Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*, 86-95.
- Anderson, R., Greene, M., & Loewen, P. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research, 34*(2), 148-165.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-574.
- Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills, CA: Sage.
- Aoki, M. (2010, September 23) Universities looking to go global. *The Japan Times*. Retrieved from <http://search.japantimes.co.jp/cgi-bin/nn20100923f1.html>
- Armor, D., Conroy-Oseguara, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading programs in selected Los Angeles minority schools*, REPORT NO. R-2007-LAUSD. Santa Monica, CA: Rand Corporation (ERIC Document Reproduction Service No. ED 130 243).
- Ashton, P. T., Olejnik, S., Crocker, L. & McAuliffe, M. (1982, April). *Measurement problems in the study of teachers' sense of efficacy*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York, NY: Longman.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*, 1175-1184.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.

- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164-180.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529.
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*(4), 676.
- Bentler, P. M. (1995). EQS structural equations program [Computer software]. Encino, CA: Multivariate Software.
- Bess, J. L. (1997). *Teaching well and liking it: Motivating faculty to teach effectively*. Baltimore, MD: Johns Hopkins University Press.
- Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring self-efficacy in pre-service elementary teachers. *School Science and Mathematics, 104*, 383-391.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Borg, S. (1998). Teachers' pedagogical system and grammar teaching: A qualitative study. *TESOL Quarterly, 32*(1), 9-38.
- Brouwers, A., & Tomic, W. (1998, July). Student disruptive behavior, perceived self-efficacy in classroom management and teacher burnout. Paper presented at the ninth European Conference on Personality, University of Surrey.
- Brutt-Griffler, J., & Samimy, K. K. (1999). Revisiting the colonial in the postcolonial: Critical praxis for nonnative English-speaking teachers in a TESOL program. *TESOL Quarterly, 33*, 413-431.
- Bureau of Labor Statistics (2012, April 10). Occupational Outlook Handbook: Postsecondary Teachers. Retrieved from <http://www.bls.gov/ooh/education-training-and-library/postsecondary-teachers.htm>
- Byrne, B. M. (2006). *Structural equation modeling with EQS* (2nd ed.). Mahwah, NJ: Erlbaum.

- Campbell, T. C., Gillaspy, J. A., & Thompson, B. (1995, January). *The factor structure of the Bem Sex-Role Inventory (BSRI): A confirmatory factor analysis*. Paper presented at the annual meeting of the Southwest Educational research Association, Dallas. (ERIC Document Reproduction Service No. ED 380 491)
- Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology, 95*(4), 821-832.
- Carver, C. S., & Scheirier, M. F. (2007). Engagement, disengagement, coping, and catastrophe. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 527-547). New York: The Guilford Press.
- Castillo, J. J. (2009). Snowball sampling. Retrieved May 12, 2011, from Experiment Resources: <http://www.experiment-resources.com/snowball-sampling.html>
- Chacón, C. T. (2002). *Teachers' sense of efficacy and selected characteristics of selected English as a foreign language Venezuelan middle school teachers*. Unpublished doctoral dissertation. The Ohio State University, Columbus, Ohio.
- Chacón, C. T. (2005). Teachers' perceived efficacy among English as a foreign language teachers in middle schools in Venezuela. *Teaching and Teacher Education, 21*, 257-272.
- Cheung, H. Y. (2008). Teacher efficacy: A comparative study of Hong Kong and Shanghai primary in-service teachers. *The Australian Educational Researcher, 35*(1), 103-123.
- Chwalisz, K., Altmaier, E. M., & Russell, D. W. (1992). Causal attributions, self-efficacy cognitions, and coping with stress. *Journal of Social and Clinical Psychology, 11*, 377-400.
- Coladarci, T., & Breton, W. (1997). Teacher efficacy, supervision, and the special education resource-room teacher. *Journal of Educational Research, 90*(4), 230-239.
- Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). City, NJ: Pearson Prentice Hall.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Davies, A. (1991). *The native speaker in applied linguistics*. Edinburgh, Scotland: Edinburgh University Press.

- Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*, 751-766.
- Denzine, G. M., Cooney, J. B., & McKenzie, R. (2005). Confirmatory factor analysis of the Teacher Efficacy Scale for prospective teachers. *British Journal of Educational Psychology, 75*, 689-708.
- Diefendorff, J. M., Richard, E. M., & Yang, J. (2008). Linking emotion regulation strategies to affective events and negative emotions at work. *Journal of Vocational behavior, 73*(3), 498-508.
- Driscoll, M. P. (2000). *Psychology of learning for instruction* (2nd ed.). Boston, MA: Allyn & Bacon.
- Dörnyei, Z. (1998). *Motivation in second and foreign language learning*. Cambridge: Cambridge University Press.
- Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching: Motivation*. New York, NY: Routledge.
- Eaton, M. J., & Dembo, M. H. (1997). Differences in the motivational beliefs of Asian American and non-Asian students. *Journal of Educational Psychology, 89*(3), 433-440.
- Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*(2), 135-172.
- Eisenberg, N., Martin, C. L., & Fabes, R. A. (1996). Gender development and gender effects. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 358-396). New York, NY: Macmillan.
- Emmer, E. (1990). *A scale for measuring teacher efficacy in classroom management and discipline*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston.
- Emmer, E., & Hickman, J. (1990). *Teacher decision making as a function of efficacy, attribution, and reasoned action*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston.
- Enochs, L. G., Scharmann, L. C., & Riggs, I. M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy and outcome expectancy. *Science Education, 79*(1), 63-75.

- Evers, W. J. G., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: a study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology*, *72*, 227-243.
- Falout, J. (2010). Strategies for teacher motivation. *The Language Teacher*, *34*(6), 27-32.
- Fisher Jr, W. P. (2007). Living capital metrics. *Rasch Measurement Transactions*, *21*(1), 1092-1095.
- Fitzpatrick, M. (2010, October 14). Japan's open-door policy hinges on an attitude shift. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/story.asp?storyCode=413828§ioncode=26>
- Freedman, S.M. & Phillips, J.S. (1985). The effects of situational performance constraints on intrinsic motivation and satisfaction: the role of perceived competence and self-determination. *Organizational Behavior and Human Decision Processes*, *35*, 397-416.
- Friedman, I. A., & Farber, B. A. (1992). Professional self-concept as a predictor of teacher burnout. *Journal of Educational Research*, *86*, 28-35.
- Gates, G.S. (2000). Teaching-related stress: The emotional management of faculty. *Review of Higher Education*, *23*(4), 469-490.
- Ghaith, G., & Shaaban, K. (1999). The relationship between perceptions of teaching concerns, teacher efficacy, and selected teacher characteristics. *Teaching and Teacher Education*, *15*(5), 487-496.
- Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*, 569-582.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Earlbaum.
- Gmelch, W., Wilke, P., & Lovrich, N. (1986). Dimensions of stress among university faculty: Faculty-analytic results from a national study. *Research in Higher Education*, *24*(3), 266-288.
- Grayson, J. L., & Alvarez, H. K. (2007). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, *24*(5), 1349-1363.
- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*, 627-643.
- Hani, Y. (2001, February 1) Colleges brace as fewer apply. *The Japan Times*. Retrieved from <http://search.japantimes.co.jp/cgi-bin/nn20010201b9.html>

- Henson, R. K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist, 37*, 137-150.
- Hooker, C., Nakamura, J., & Csikszentmihalyi, M. (2003). The group as mentor: Social capital and the systems model of creativity. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 225-244). New York, NY: Oxford University Press.
- Horwitz, E. K. (1996). Even teachers get the blues: Recognizing and alleviating language teachers' feelings of foreign language anxiety. *Foreign Language Annals, 29*(3), 365-372.
- Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal, 93*, 356-372.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.
- Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*(1), 1-35.
- Imants, J. G., & De Brabander, C. J. (1996). Teachers' and principals' sense of efficacy in elementary schools. *Teaching and Teacher Education, 12*(2), 179-195.
- Joreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). London, England: Sage.
- Karimvand, P., N. (2011). The Nexus between Iranian EFL Teachers' self-efficacy, teaching experience and gender. *English Language Teaching, 4*, 171-183.
- Kieschke, U., & Schaarschmidt, U. (2008). Professional commitment and health among teachers in Germany: A typological approach. *Learning and Instruction, 18*(5), 429-437.
- Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y., & Georgiou, T. (2009). Exploring the validity of the Teachers' Self-Efficacy Scale in five countries. *Contemporary Educational Psychology, 34*, 67-76.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

- Kooij, D., De Lange, A., Jansen, P., & Dijkers, J. (2008). Older workers' motivation to continue to work: Five meanings of age: A conceptual review. *Journal of Managerial Psychology*, 23(4), 364-394.
- Labone, E. (2004). Teacher efficacy: Maturing the construct through research in alternative paradigms. *Teaching and Teacher Education*, 20, 341-359.
- Law, G. (1995). Ideologies of English language education in Japan. *JALT Journal*, 17(2), 213-224.
- Leiter, M. P. (1992). Burnout as a crisis in self-efficacy: Conceptual and practical implications. *Work and Stress*, 6(2), 107-115.
- Lent, R. W., & Hackett, G. (1987). Career self-efficacy: Empirical status and future directions. *Journal of Vocational Behavior*, 30(3), 347-382.
- Levin, J. D., Culkin, J., & Perrotto, R. S. (2001). *Introduction to chemical dependency counseling*. Northvale, NJ: Aronson.
- Lin, S. S., & Tsai, C. C. (1999). Teaching Efficacy along the Development of Teaching Expertise among Science and Math Teachers in Taiwan. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Boston, MA.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3, 85-106.
- Linacre, J. M. (2006). Data variance explained by Rasch measures. *Rasch Measurement Transactions*, 20(1), 1045.
- Linacre, J. M. (2009). A user's guide to WINSTEPS: Rasch-model computer program. Chicago, IL: MESA.
- Linacre, J. M., & Wright, B. D. (2009). WINSTEPS: Multiple-choice, rating scale, and partial credit Rasch analysis [Computer software]. Chicago, IL: MESA.
- Liu, J. (2009). Complexities and Challenges in Training Nonnative English Speaking Teachers: State-of-the-Art. CamTESOL conference materials. Retrieved on July 2, 2013 at http://www.camtesol.org/2009conference/Documents/Jun_Liu_09.ppt/
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705-718.

- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago, IL: Chicago University Press.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224.
- Maslach, C. (1993). Burnout: A multidimensional perspective. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.). *Professional burnout: Recent developments in theory and research* (pp. 19-32). Washington, DC: Taylor and Francis.
- Mau, W. C. J. (2004). Cultural Dimensions of Career Decision-Making Difficulties. *The Career Development Quarterly*, 53(1), 67-77.
- McConnell, D. L. (2000). *Importing diversity: Inside Japan's JET Program*. Berkeley, CA: University of California Press.
- McCurry, J. (2011, March 8) Japan launches primary push to teach English. *The Guardian*. Retrieved from <http://www.guardian.co.uk/education/2011/mar/08/japan-launches-primary-english-push>
- McLaughlin, M. W., & Marsh, D. D. (1978). Staff development and school change. *Teachers College Record*, 80, 70-94.
- McNeil, D. (2010, September 10). Japan's Globalization Project Stalls as Some Criticize Focus on Elite Universities. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/A-Slow-Start-for-Japans/124346/>
- Medgyes, P. (1994). *The non-native teacher*. London, England: Macmillan.
- Meijer, C., & Foster, S. (1988). The effect of teacher self-efficacy on referral chance. *Journal of Special Education*, 22, 378-385.
- Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 81, 247-258.
- Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals*, 39(2), 276-295.
- Mills, N. (2011). Teaching assistants' self-efficacy in teaching literature: Sources, personal assessments, and consequences. *The Modern Language Journal*, 95, 61-80.

- Milner, H. R. (2001). A qualitative investigation of teachers' planning and efficacy for student engagement. *Unpublished Ph.D. dissertation*. The Ohio State University, Columbus, OH.
- Milner, H. R. (2002). A case study of an experienced teachers' self efficacy and persistence through crisis situations: Theoretical and practical considerations. *The High School Journal*, 86, 28-35.
- Milner, H. R., & Hoy, A. W. (2003). A case study of an African American Teacher's self-efficacy, stereotype threat, and persistence. *Teaching and Teacher Education*, 19, 263-276.
- Mowday, R. T. & Nam, S. H. (1997). Motivating more effective university teaching: The implications of expectancy theory. In J. L. Bess (Ed.), *Teaching well and liking it* (pp. 110-124). Baltimore, MD: Johns Hopkins University Press.
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology*, 38(1), 30-38.
- Murray, A. (2013). Teacher burnout in Japanese higher education. *The Language Teacher*, 37(4), 51-55.
- Mushayikwa, E., & Lubben, F. (2009). Self-directed professional development—Hope for teachers working in deprived environments? *Teaching and Teacher Education*, 25(3), 375-382.
- Neisser, U., & Jopling, D. (Eds.). (1997). *The Conceptual Self in Context: Culture, experience, self-understanding*. New York: Cambridge University Press.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, 37(4), 589-613.
- Olsen, D. (1993). Work satisfaction and stress in the first and third year of academic appointment. *Journal of Higher Education*, 64, 453-471.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66, 543-578.
- Pajares, M. F. (2002). Overview of social cognitive theory and of self-efficacy. Retrieved January 10, 2011, from <http://www.emory.edu/EDUCATION/mfp/eff.html>
- Patton, M. (2001). *Qualitative research & evaluation methods*. Thousand Oaks, CA: Sage.

- Phillips, J. S., & Freedman, S. M. (1984). Situational performance constraints and task characteristics: Their relationship to motivation and satisfaction. *Journal of Management, 10*, 321-331.
- Poppleton, P., & Riseborough, G. (1990). A profession in transition: Educational policy and secondary school teaching in England in the 1980s. *Comparative Education, 26*, 211-226.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Denmark Paedagogiske Institut.
- Riggs, I. (1995). *The characteristics of high and low efficacy elementary teachers*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco.
- Riggs, I., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy instrument. *Science Education, 74*, 625-637.
- Riggs, I., & Jesunathadas, J. (1993). *Preparing elementary teachers for effective science teaching in diverse settings*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.
- Rivers, D. J. (2013). Labor contract law amendments: Recruitment indicative of change? *The Language Teacher, 37*(1), 68-71.
- Roberts, J. K., & Henson, R. K. (2000). *Self-Efficacy Teaching and Knowledge Instrument for Science Teachers (SETAKIST): A proposal for a new efficacy instrument*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, KY.
- Roberts, J. K., & Henson, R. K. (2001). *A confirmatory factor analysis of a new measure of teacher efficacy: Ohio State Teacher Efficacy Scale*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcome. *Journal of Educational Research, 74*, 185-190.
- Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 95*, 534-562.
- Ross, J. A. (1994). The impact of an in-service to promote cooperative learning on the stability of teacher efficacy. *Teaching and Teacher Education, 10*, 381-394.

- Ross, J. A. (1995). Strategies for enhancing teachers' beliefs in their effectiveness: Research on a school improvement hypothesis. *Teachers' College Record*, 97(2), 227-251.
- Ross, J. A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Advances in research on teaching*, Vol. 7 (pp. 49-73). Greenwich, CT: JAI Press.
- Rotter, J. B. (1954). *Social learning and clinical psychology*. New York, NY: Prentice-Hall.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, 1-28.
- Sawa, T. (2004, November 1) Universities lack will to reform. *The Japan Times*. Retrieved from <http://search.japantimes.co.jp/cgi-bin/ea20041101ts.html>
- Sax, L. J. (1996). The American college teacher: National norms for the 1995-96 HERI faculty study. *ERIC Document Reproduction Service*, No. ed399863.
- Schaufeli, W. B., Maslach, C., & Marek, T. (1993). *Professional burnout: Recent developments in theory and research*. Washington, DC: Taylor and Francis.
- Scholz, U. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European journal of psychological assessment*, 18(3), 242-251.
- Schunk, D. H. (1989). Self-efficacy and cognitive skill learning. In C. Ames & R. Ames (Eds.), *Research on motivation in education: Vol. 3. Goals and cognitions* (pp. 13-44). San Diego, CA: Academic.
- Schwarzer, R., & Born, A. (1997). Optimistic self-beliefs: Assessment of general perceived self-efficacy in thirteen cultures. *World Psychology*, 3(1-2), 177-190.
- Sergiovanni, T. (1967). Factors which affect satisfaction and dissatisfaction of teachers. *Journal of Educational Administration*, 5(1), 66-82.
- Shim, J. W. (2001). *The efficacy beliefs of Korean teachers of English as a foreign language*. Unpublished doctoral dissertation. The Ohio State University, Columbus, Ohio.
- Shim, J. W. (2003). Teacher efficacy beliefs and language skills of pre-service elementary teachers of English. *Studies on English Language and Literature*, 29, 231-244.

- Shim, J. W. (2006). Validation of dimensions of teacher efficacy beliefs specific to non-native Korean teachers of English. *Studies on English Language and Literature*, 32, 207-224.
- Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, 26, 1059-1069.
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effect of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571-581.
- Snyder, C. R. (1994). *The psychology of hope: You can get there from here*. New York, NY: Free Press.
- Snyder, C. R., Cheavens, J., & Sympson, S. C. (1997). Hope: An individual motive for social commerce. *Group Dynamics: Theory, Research, and Practice*, 1(2), 107.
- Soodak, L., & Podell, D. (1996). Teaching efficacy: Toward the understanding of a multi-faceted construct. *Teaching and Teacher Education*, 12(4), 401-412.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Tabachnick, B. G., & Fidell, L. S., (2007). *Using multivariate analysis* (5th ed.). Boston, MA: Allyn and Bacon.
- Tang, C. (1997). On the power and status of nonnative ESL teachers. *TESOL Quarterly*, 31(3), 577-580.
- Tesch, R. (1990). *Qualitative research: Analysis types and software tools*. New York, NY: Falmer.
- Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality*, 69(6), 907-924.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure *Review of Educational Research*, 68, 202-248.
- Van Dierendonck, D., Schaufeli, W. B., & Sixma, H. (1994). Burnout among general practitioners: A perspective from equity theory. *Journal of Social and Clinical Psychology*, 13, 86-100.

- Vandenberghe, R., & Huberman, A. M. (1999). *Understanding and preventing teacher burnout: A sourcebook of international research and practice*. New York, NY: Cambridge University Press.
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research, 54*(2), 143-178.
- Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*, 231-251.
- Waiters, J. J., & Ginns, I. S. (1995). *Origins of and changes in preservice teachers' science teaching efficacy*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco.
- Walker, C. J., & Symons, C. (1997). The meaning of human motivation. In Bess, J. L. (Ed.). *Teaching well and liking it: Motivating faculty to teach effectively* (pp. 3-18). Baltimore, MD: The Johns Hopkins University Press.
- White, S. A., Eguchi, M., Kawanaka, J., & Henneberry, S. W. (2005). Addressing declining academic skills among Japanese college and university students: Are study support services the answer? *Shimane Journal of Policy Studies, 9*, 35-52.
- Whitley, B. E. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in Human Behavior, 13*(1), 1-22.
- Williams, M., & Burden R. L. (1997). *Psychology for language teachers: A social constructivist approach*. Cambridge, England: Cambridge University Press.
- Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology, 99*, 181-193.
- Woolfolk Hoy, A., & Burke Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21*, 343-356.
- Woolfolk Hoy, A., Davis, H., & Pape, S. (2006). Teacher knowledge, beliefs, and thinking. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 709-725). Mahwah, NJ: Erlbaum.
- Woolfolk Hoy, A., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of Educational Psychology, 82*, 81-91.
- Woolfolk, A. E., Rosoff, B., & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6*, 137-148.

Wright, B. D. (2000). Conventional factor analysis vs. Rasch residual factor analysis. *Rasch Measurement Transactions*, 14(2), 753.

Zeldin, A. L., & Pajares, F. (2000). Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal*, 37(1), 215-246.

APPENDIX A
ONLINE CONSENT FORM

Temple University Japan
Online Consent Form

Dear Colleagues,

Hello! My name is Max Praver and I'm a current doctoral student at Temple University Japan. I am writing to ask your assistance in a study of Japanese university English language teachers.

Purpose of the Study: This study explores what English language educators face at Japanese universities.

Confidentiality: Your individual privacy will be maintained in all written and published works, as well as presentations resulting from this study. The data gathered will be used for research purposes only and your answers will be released only as summaries, in which no individual's answers can be identified. The confidentiality of records identifying the participant and their institution(s) will be maintained by the use of pseudonyms and identification numbers.

Voluntary Participation: Participation is voluntary. You have the right to contact me and withdraw at any time. You also have the right to instruct me at any time not to use your data.

Procedure: If you agree to participate, the survey will take approximately 10 minutes to finish. Please complete it in one sitting. There are 2 parts. In Part 1, please complete the background questionnaire. In Part 2, please indicate your answers about your teaching situation and university activities on the 6-point scale. When you have finished the survey, please put both parts 1 and 2 in the self-addressed stamped envelope provided and drop it in any post office box.

Contact Information and Questions: Should you have any questions at all concerning the survey or wish to receive a copy of the final results, please don't hesitate to contact me at xxxxxxxx@xxx.com or 090-XXXX-XXXX.

Thank you once again. Your help is very much appreciated.

Sincerely,
Max Praver

APPENDIX B

BACKGROUND QUESTIONNAIRE

Please complete the following background questionnaire. As stated on the consent form, all information will be kept strictly confidential. Thank you.

1. Gender

Male Female

2. Age

20-29 30-39 40-49 50-59 60-69

3. What is your nationality? _____

4. Do you consider English to be your first language?

Yes No

Questions 5-8 are for people who answered 'No' to the previous question

5. If English is not your first language, how and where did you learn English?

6. If English is not your first language, what is the total amount of time you've spent in an English speaking country?

7. What was the purpose of that time abroad? (e.g., study, work, travel, etc.)

8. If you can recall your most recent TOEFL or TOEIC score, please write it below.

9. What is your highest completed educational degree?

Bachelors Masters Doctorate

10. What was that degree in? (e.g., Applied Linguistics, TESOL, British Lit, etc.)

11. How many years have you been teaching English? _____

12. Which kind of students do you primarily teach?

Undergraduate Graduate Other _____

13. Are your students primarily English majors?

Yes No

14. Which of the following best describes your work status now?

- a. Tenured
- b. Non-tenured term-limited contract
- c. Non-tenured unlimited contract (infinitely renewable)
- d. Part-time

15. How many koma per week do you teach? _____

16. What is the average number of students in the classes you teach? If it varies greatly, please briefly explain. _____

17. Please indicate your title

- Lecturer / Instructor
- Assistant Professor
- Associate Professor
- Professor
- Other _____

Questions 18-20 are for non-tenured full-time contract teachers only

18. What is the length of your current contract? _____ year(s)

19. What year of the contract are you in now? _____

20. Is your contract renewable? If yes, how many times? _____

21. What is your self-evaluated Japanese ability?

Beginner Intermediate Advanced Fluent/1st language

22. Please write the name of the institution(s) where you currently work.

23. Do you feel Japanese language skills are necessary when communicating with your superiors at your institution(s)?

Yes No

Please briefly explain

24. In the event that you would be willing to participate in a follow-up interview, please indicate your agreement to do so below. The purpose of the interview is to help me gain a deeper understanding about the topics being asked in the questionnaire and the survey results. The interview will be conducted at your convenience and its estimated time is about 60 to 90 minutes. An interview participant will be provided an opportunity to openly talk with me about the issues that he/she faces in teaching English at Japanese universities. Interviews will be audio recorded and all data will be kept strictly confidential. Also, each interview participant will receive a free drink and a 1,000-yen gift certificate as an expression of my gratitude.

I would _____/would not _____ like to participate in an interview with Max Praver at a later time.

I can be reached at the following contact information:

Name: _____

E-mail address: _____

Thank you very much!
Part 2 of the survey begins on the next page.

APPENDIX C

JAPANESE UNIVERSITY LANGUAGE TEACHER EFFICACY BELIEFS SCALE

Please indicate your answers on the scale provided for each of the following statements. Your answers will be kept confidential.

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
1. I can communicate in a meaningful way with my students.	(1)	(2)	(3)	(4)	(5)	(6)
2. I can respond effectively to English language questions from my students.	(1)	(2)	(3)	(4)	(5)	(6)
3. I can control disruptive behavior in my English classroom.	(1)	(2)	(3)	(4)	(5)	(6)
4. I can communicate my opinions effectively to my superiors.	(1)	(2)	(3)	(4)	(5)	(6)
5. I can help my students think critically in English.	(1)	(2)	(3)	(4)	(5)	(6)
6. I can gauge my students' comprehension of what I have just taught.	(1)	(2)	(3)	(4)	(5)	(6)
7. I can make my expectations clear about appropriate student behavior.	(1)	(2)	(3)	(4)	(5)	(6)
8. I can convince my superiors to listen to my suggestions.	(1)	(2)	(3)	(4)	(5)	(6)
9. I can motivate most of my students to become more interested in English.	(1)	(2)	(3)	(4)	(5)	(6)
10. I can help students become more independent English learners.	(1)	(2)	(3)	(4)	(5)	(6)
11. I can convince my superiors to act on my suggestions.	(1)	(2)	(3)	(4)	(5)	(6)

- | | | | | | | |
|--|-----|-----|-----|-----|-----|-----|
| 12. I can create a learning friendly environment in my classroom. | (1) | (2) | (3) | (4) | (5) | (6) |
| 13. I can deal effectively with defiant students. | (1) | (2) | (3) | (4) | (5) | (6) |
| 14. I can provide appropriate support for my least proficient English language students. | (1) | (2) | (3) | (4) | (5) | (6) |
| 15. I can use a variety of language assessment strategies in my classes. | (1) | (2) | (3) | (4) | (5) | (6) |
| 16. I can get my students to believe they can succeed in English class. | (1) | (2) | (3) | (4) | (5) | (6) |
| 17. I can establish effective routines in my classes. | (1) | (2) | (3) | (4) | (5) | (6) |
| 18. I can provide appropriate challenges for my most proficient English language students. | (1) | (2) | (3) | (4) | (5) | (6) |
| 19. I can get my students to follow classroom rules. | (1) | (2) | (3) | (4) | (5) | (6) |
| 20. I can avoid being adversely affected by school bureaucracy. | (1) | (2) | (3) | (4) | (5) | (6) |
| 21. I can convince my superiors to take my suggestions seriously. | (1) | (2) | (3) | (4) | (5) | (6) |
| 22. I can get most of my students to value learning English. | (1) | (2) | (3) | (4) | (5) | (6) |
| 23. I can get my superiors to support decisions I make regarding English teaching. | (1) | (2) | (3) | (4) | (5) | (6) |
| 24. I can foster student creativity. | (1) | (2) | (3) | (4) | (5) | (6) |
| 25. I can put together level appropriate learning activities for my students. | (1) | (2) | (3) | (4) | (5) | (6) |
| 26. I can adjust my lessons to different levels for individual students. | (1) | (2) | (3) | (4) | (5) | (6) |
| 27. I can establish my own classroom management system. | (1) | (2) | (3) | (4) | (5) | (6) |

28. I can persuade my superiors to accommodate my teaching needs. (1) (2) (3) (4) (5) (6)
29. I can provide alternative explanations when students are confused. (1) (2) (3) (4) (5) (6)
30. I can prevent problem students from ruining classes. (1) (2) (3) (4) (5) (6)
31. I can greatly improve the overall English skills of my students. (1) (2) (3) (4) (5) (6)
32. I can convince my superiors to have greater vision regarding English language teaching. (1) (2) (3) (4) (5) (6)

Thank you very much for participating in this study. Should you have any questions, please don't hesitate to contact me at: xxxx@xxxx.com

Sincerely,
Max Praver

APPENDIX D

THE 42 SELF-EFFICACY ITEMS (INCLUDING THE 24 ITEMS ADAPTED FROM THE TSES) USED IN THE PILOT STUDY

Efficacy in Student Engagement (SE)

1. I can communicate in a meaningful way with my students.
2. I can help my students think critically in English.
3. I can motivate students to become more interested in English.
4. I can help students become more independent English learners.
5. I can get my students to believe they can succeed in English class.
6. I can get my students to value learning English.
7. I can foster student creativity.
8. I can improve the overall English skills of my students.

Efficacy in Instructional Strategies (IS)

1. I can respond effectively to English language questions from my students.
2. I can gauge my students' comprehension of what I have just taught.
3. I can keep activities running smoothly for the entire lesson.
4. I can teach an English language concept better after having taught it once before.
5. I can improvise when necessary in my English language classes.
6. I can put together level appropriate learning activities for my students.
7. I can adjust my lessons to different levels for individual students.
8. I can gauge my students' English progress over the semester.
9. I can use a variety of language assessment strategies in my classes.
10. I can implement alternative teaching strategies in my classroom to accommodate the various levels of students.
11. I can provide alternative explanations when students are confused.
12. I can provide appropriate challenges for my most proficient English language students.
13. I can provide appropriate support for my least proficient English language students.

Efficacy in Classroom Management (CM)

1. I can control disruptive behavior in my English classroom.
2. I can make my expectations clear about appropriate student behavior.
3. I can establish effective routines in my classes.
4. I can get my students to follow classroom rules.
5. I can establish my own classroom management system.
6. I can prevent problem students from ruining classes.
7. I can deal effectively with uncooperative students.

8. I can deal effectively with defiant students.
9. I can create a learning friendly environment in my classroom.

Efficacy in Dealing with Superiors (DS)

1. I can communicate my opinions effectively to my superiors.
2. I can convince my superiors to listen to my suggestions.
3. I can convince my superiors to act on my suggestions.
4. I can deal with rules created by my superiors.
5. I can choose which classes I want to teach.
6. I can avoid being adversely affected by school bureaucracy.
7. I can persuade my superiors to accommodate my teaching needs.
8. I can fail students without approval.
9. I can convince my superiors to have greater vision regarding language teaching.
10. I can get my superiors to support decisions I make regarding English teaching.
11. I can get my superiors to accommodate my desired teaching schedule.
12. I can convince my superiors to take my suggestions seriously.

APPENDIX E

THE 32 SELF-EFFICACY ITEMS

Efficacy in Student Engagement (SE)

1. I can communicate in a meaningful way with my students.
2. I can help my students think critically in English.
3. I can motivate most of my students to become more interested in English.
4. I can help students become more independent English learners.
5. I can get my students to believe they can succeed in English class.
6. I can get most of my students to value learning English.
7. I can foster student creativity.
8. I can greatly improve the overall English skills of my students.

Efficacy in Instructional Strategies (IS)

1. I can respond effectively to English language questions from my students.
2. I can gauge my students' comprehension of what I have just taught.
3. I can provide appropriate support for my least proficient English language students.
4. I can use a variety of language assessment strategies in my classes.
5. I can provide appropriate challenges for my most proficient English language students.
6. I can put together level appropriate learning activities for my students.
7. I can adjust my lessons to different levels for individual students.
8. I can provide alternative explanations when students are confused.

Efficacy in Classroom Management (CM)

1. I can control disruptive behavior in my English classroom.
2. I can make my expectations clear about appropriate student behavior.
3. I can create a learning friendly environment in my classroom.
4. I can deal effectively with defiant students.
5. I can establish effective routines in my classes.
6. I can get my students to follow classroom rules.
7. I can establish my own classroom management system.
8. I can prevent problem students from ruining classes.

Efficacy in Dealing with Superiors (DS)

1. I can communicate my opinions effectively to my superiors.
2. I can convince my superiors to listen to my suggestions.
3. I can convince my superiors to act on my suggestions.
4. I can avoid being adversely affected by school bureaucracy.
5. I can convince my superiors to take my suggestions seriously.
6. I can get my superiors to support decisions I make regarding English teaching.
7. I can persuade my superiors to accommodate my teaching needs.
8. I can convince my superiors to have greater vision regarding language teaching.

APPENDIX F

THE FINAL 28 POST-INSTRUMENT VALIDATION SELF-EFFICACY ITEMS USED IN THE MAIN ANALYSIS

Efficacy in Student Engagement (SE)

2. I can help my students think critically in English.
3. I can motivate most of my students to become more interested in English.
4. I can help students become more independent English learners.
5. I can get my students to believe they can succeed in English class.
6. I can get most of my students to value learning English.
8. I can greatly improve the overall English skills of my students.

Efficacy in Instructional Strategies (IS)

1. I can respond effectively to English language questions from my students.
2. I can gauge my students' comprehension of what I have just taught.
3. I can provide appropriate support for my least proficient English language students.
4. I can use a variety of language assessment strategies in my classes.
5. I can provide appropriate challenges for my most proficient English language students.
6. I can put together level appropriate learning activities for my students.
7. I can adjust my lessons to different levels for individual students.
8. I can provide alternative explanations when students are confused.

Efficacy in Classroom Management (CM)

1. I can control disruptive behavior in my English classroom.
2. I can make my expectations clear about appropriate student behavior.
3. I can create a learning friendly environment in my classroom.
4. I can deal effectively with defiant students.
5. I can establish effective routines in my classes.
6. I can get my students to follow classroom rules.
7. I can establish my own classroom management system.
8. I can prevent problem students from ruining classes.

Efficacy in Dealing with Superiors (DS)

2. I can convince my superiors to listen to my suggestions.
3. I can convince my superiors to act on my suggestions.
5. I can convince my superiors to take my suggestions seriously.
6. I can get my superiors to support decisions I make regarding English teaching.
7. I can persuade my superiors to accommodate my teaching needs.
8. I can convince my superiors to have greater vision regarding language teaching.

APPENDIX G

INTERVIEW CONSENT FORM

Temple University Japan Informed Consent Form

The following is a consent form for my doctoral research project at Temple University Japan.

Consent for Participation in Interview Research

I volunteer to participate in a research project conducted by Max Praver from Temple University Japan. I understand that the project is designed to gather information about what English language educators face at Japanese universities. I will be one of approximately 20 people being interviewed for this research and I will receive a 1000 Yen gift certificate for my time.

1. My participation in this project is voluntary. I may withdraw and discontinue participation at any time without penalty or need for explanation. If I decline to participate or withdraw from the study, none of my data will be used.
2. Participating in this study will be a chance for me to discuss my experiences of language teaching and will help the researcher learn more about the challenges of working as an English teacher at Japanese universities. I understand that most interviewees will find the discussion interesting and thought-provoking. If, however, I feel uncomfortable in any way during the interview session, I have the right to decline to answer any question or to end the interview.
3. Participation involves being interviewed by Max Praver from Temple University. The interview will last approximately 60-90 minutes. The researcher will digitally record the interview. I will not be asked to state my name or institution on the recording. The recordings and the transcripts of the recordings, without my name, will be kept until the research is complete.
4. I understand that the researcher will not identify me by name in any reports using information obtained from this interview, and that my confidentiality as a participant in this study will remain secure. Subsequent uses of records and data will be subject to standard data use policies, which protect the anonymity of individuals and institutions.
5. Faculty and administrators from my institution will neither be present at the interview nor have access to recordings, raw notes, or transcripts. This precaution will prevent my individual comments from having any negative repercussions.
6. I have read and understand the explanation provided to me. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this study.
7. I have been given a copy of this consent form.

_____ My Signature _____ Date

_____ My Printed Name _____ Signature of the Investigator

For further information or questions not addressed by the researcher please contact:
Dr. David Beglar, Academic Coordinator, Temple University Japan [dbeglar@me.com]

APPENDIX H

GUIDING INTERVIEW QUESTIONS

1. What attracts you to teaching English now? What attracted you originally at the beginning?
2. Have you ever worked in any other professions? If yes, what professions and why the change to teaching English?
3. Tell me about any formal pre- or in-service professional training experiences.
4. What parts of your job are easiest? Why?
5. What's the hardest part about being a university English teacher in Japan? Why? (If relevant) How does this compare with the hardest part of professions you have worked in?
6. What's the best part of your job? Why?
7. What's the worst part of your job? Why?
8. What aspects of your job (if any) make it difficult for you to do it well? In class? Out of class? Any other impediments? Why? How do you deal with such challenges?
9. What aspects of your teaching situation would you like to improve? Why?
10. What sort of things could help you with these improvements? (Note: do not mention the following possibilities unless the interviewee gets 'stuck') More external support (financial, professional development, etc...)? More knowledge? Other?
11. What kind of teaching methods/activities/materials do you use and what are the major objectives of your classes? And are these decided by you and communicated for approval as part of the course plan? Decided by you but not reviewed by anyone? Decided in collaboration or with a committee?
12. What have been your major sources of information about teaching English? [Sources of efficacy information]
13. Could you briefly describe your experiences teaching English at Japanese universities? [Mastery experiences]
14. Do/have you ever received feedback from others about your teaching? What types of feedback? (observations, comments, student evaluations, etc...) Who, if anyone, has provided you encouragement/and or strategies for overcoming obstacles in teaching English? [Verbal persuasion]
15. Do/have other teachers at all influence your teaching? How could you imagine yourself influenced by others (i.e., observing others teach, media/society images, mentors of successful teachers, comparisons to others)? [Vicarious experiences]
16. How do you feel when you are teaching a successful/smooth English class/lesson/activity? [Physiological and Emotional Cues]
17. How do you feel when you are teaching an unsuccessful/rocky English class/lesson/activity? [Physiological and Emotional Cues]