Measuring English Language Self-Efficacy: Psychometric Properties and Use

Do-Hong Kim
College of Education, Wayne State University, USA

Chuang Wang (Corresponding Author)
Faculty of Education, University of Macau, China
Email: wangc@um.edu.mo

Ting Sun
Cato College of Education, University of North Carolina at Charlotte, USA

Lan Liu
Faculty of Hospitality and Tourism Management, Macau University of Science and Technology, China

Abstract

This article provides an overview of existing instruments measuring self-efficacy for English language learning in both first and second language acquisition fields and their reliability and validity evidence. It also describes the development and use of the Questionnaire of English Language Self-Efficacy (*QESE*) scale, designed specifically for English language learners (ELLs), and presents an overview of the research findings from empirical studies related to its psychometric properties. A growing body of literature has begun to document encouraging evidence of ELL students' self-efficacy belief measures and the utility of the *QESE* in particular. The information pertaining to the *QESE* is quite encouraging from measurement perspectives and fills the gap in the literature by providing a reliable and valid instrument to measure ELLs' self-efficacy in various cultures. This paper concludes with evidence for internal consistency, test-retest reliability, structural, generalizability, and external aspects of the construct validity of the QESE. This paper contributes to the growing interest in these skills by reviewing the measures of self-efficacy in the field of second-language acquisition and the findings of empirical research on the development and use of a self-efficacy scale for ELLs.
Keywords: English language learners; Second language acquisition; Self-efficacy; Measurement; Psychometric properties.

1. Introduction

Self-efficacy refers to a person's belief in successfully completing a task based upon his/her self-assessment of competence (Bandura, 2006). Self-efficacy is very important in learning because it is associated with one's motivation, affect, and behaviors (Bandura, 2006). Self-efficacious students use more cognitive and metacognitive strategies and persist longer in the face of adversity compared to their less efficacious counterparts (Pajares, 2009). The positive relationship between self-efficacy and academic achievement in first language English, writing, and mathematics have been documented for decades (e.g., (Boekaerts and Cascallar, 2006; Pajares and Valiante, 2002; Pape and Wang, 2003; Zimmerman and Cleary, 2006), but empirical studies of English language learners' self-efficacy have just recently caught researcher's attention. (Where is your comment on these citations?)

2. English Language Learners' Self-Efficacy

Extensive studies on self-efficacy of English language learners (ELL) have been conducted in many countries such as China (Huang *et al.*, 2015; Wang *et al.*, 2012; Wang and Bai, 2017), Germany (Wang *et al.*, 2013a), Iran (Zabihi, 2018), Korea (Kim *et al.*, 2015; Wang *et al.*, 2013b), Norway (Solheim, 2011), Spain (Ivars *et al.*, 2014), and the United States (Zuo and Wang, 2016). Self-efficacy beliefs were noted to have a strong and positive relationship to English language proficiency and the use of strategies to study the language. For example, a recent study in Hong Kong examined the relationship between self-efficacy beliefs and English language achievement among 1092 Chinese students and reported that self-efficacy strongly predicted English language proficiency (Bai *et al.*, 2018).

Various domain-specific areas have been explored with ELL's self-efficacy beliefs. Specifically, Rahimi and Abedini (2009) examined self-efficacy in listening and Sarshar and Oroji (2016) investigated self-efficacy in speaking. A lot more scholars stayed in the areas of reading (Amogne, 2008; Shang, 2010; Smith *et al.*, 2012; Solheim, 2011) and writing (Amogne, 2008; Woodrow, 2011; Zhang and Guo, 2012), and some researchers studied self-efficacy in interpretation proficiency (e.g., (Ivars *et al.*, 2014). Solheim (2011) noted a positive relationship between reading self-efficacy and reading comprehension performance with fifth-grade-students while controlling for listening comprehension skills, reading ability, and nonverbal ability. In another study with college students in Spain, self-efficacy was found to contribute significantly to the prediction of the accuracy of interpretation between Spanish and English languages (Ivars *et al.*, 2014). Another study noted significant predictive as well as mediation roles of self-efficacy in the relationship between writing performance and anxiety (Woodrow, 2011).

Consistent with Bandura's sources on the self-efficacy theory, previous studies suggested that sources on the development of self-efficacy include self-awareness, past performance, task familiarity and difficulty level, social persuasion, physiological or emotional state, and interest (Wang and Pape, 2007; Zuo and Wang, 2016). Not all previous studies, however, were successful in providing evidence of a positive relationship between self-efficacy and English language proficiency. For example, Anastasiou and Michail (2013) found that the self-efficacy beliefs of adult students in Second-Chance Schools with learning disabilities after an 8-month instruction did not make an improvement in their writing performance, but their writing self-efficacy increased significantly. The researchers further noted that the discordance could be attributed to the literacy and writing practices such as performance ambiguity, indefinite writing aims, and problems in writing performance feedback followed in the instruction (Anastasiou and Michail, 2013). Corkette *et al.* (2011), also noted non-significant relationships between six graders' literacy ability (i.e., reading and writing abilities) and their self-efficacy. The researchers indicated that this non-significant result may be due to the inaccurate perceptions of self-efficacy rated by young students, which would become more accurate with aging (Corkette *et al.*, 2011).

2.1. Measurement of ELL's Self-Efficacy

Since self-efficacy is an important construct in the teaching and learning of English as a foreign language, the measurement of ELL's self-efficacy becomes pivotal. According to Messik (1995), the interpretation and use of the scores is the most important aspect of construct validity. The measurement of language learners' self-efficacy beliefs was carried out in the 20th century but with many difficulties. Without a sound theory and foundations of empirical studies, some scholars developed instruments to measure self-efficacy but mistakenly measured anxiety, performance expectancy, and language aptitude (e.g., (Tremblay and Gardner, 1995; Yang, 1999). Relying on inadequate operational definitions of self-efficacy is a threat to construct validity (Bandura, 1997). With a lack of clear understanding of the construct, tasks used to measure the construct, and the characteristics of contexts under which the tasks are performed are common causes for inadequate measurement of self-efficacy (Bong, 2006). Confusions of the definition of self-efficacy with other similar constructs such as self-confidence, self-esteem, and locus of control are also common mistakes in the measurement of self-efficacy beliefs (Bong, 2006).

The most popular instrument to measure students' self-efficacy is the "Self-efficacy for Learning and Performance" subscale of the Motivated Strategies for Learning Questionnaire (Wang *et al.*, 2013a). This subscale consists of eight items designed to measure student beliefs of how well they can succeed in an undergraduate course. Although this subscale is a valid measure of student self-efficacy (Vogt, 2003), the scale is best used in the context of a particular course and does not address the tasks commonly encountered in the language learning context. According to Bandura (2006), the accurate measurement of self-efficacy has to be tailored to the specific domain. Therefore, researchers in the field of English language learning developed a few instruments to measure student self-efficacy to perform specific language learning tasks in listening (Chen, 2007), reading (Baldwin *et al.*, 1980; Li and Wang, 2010; Shell *et al.*, 1995), and writing (Graham and Harris, 1989; Pajares *et al.*, 2001; Pajares, 2007; Prat-Sala and Redford, 2010; Shell *et al.*, 1989; Zimmerman and Bandura, 1994). Apart from these scholars who tried to measure specific English language tasks, some scholars developed instruments to measure English language acquisition in general (e.g., (Mills, 2004; Wang, 2004; Wong, 2005).

The purpose of this paper is to provide an overview of existing instruments measuring self-efficacy for English language learning in both first and second language acquisition fields and their reliability and validity evidence. This paper also describes the development and use of the Questionnaire of English Language Self-Efficacy (*QESE*) scale, designed specifically for ELLs, and presents an overview of the research findings from empirical studies related to its psychometric properties. This information is especially timely and pertinent for ELL teachers and researchers given the non-cognitive skills' central roles in second language acquisition.

Toward this end, we provide some recommended practices in English language classrooms and directions for future research with respect to ELL's self-efficacy. This paper contributes to the growing interest in these skills by reviewing the measures of self-efficacy in the field of second-language acquisition and the findings of empirical research on the development and use of a self-efficacy scale for ELLs.

3. Methods
3.1. Literature Search

A systematic review of the scientific literature was undertaken to locate all empirical articles examining self-efficacy for English language learning in both first and second language acquisition fields because it is believed that an instrument to measure this contract would have been reported in these articles. Articles were located through computerized journal databases (e.g., PsychInfo, Google Scholar, & JSTOR), using combinations of the following keywords: self-efficacy, second language acquisition, English language learners, measurement, psychometric. The inclusion criteria included: (a) empirical research; (b) peer-reviewed; (c) published in 2004-2019 and (d) in a formal language learning setting. With these key words and inclusion criteria, a search of the PsycInfo, Google Scholar, JSTOR, ERIC, and Web of Science databases returned 89 original research studies that were eligible for review. Articles dated earlier than 2004 were found in the citations of the articles located using the inclusion criteria.

Empirical studies on validity and reliability evidence of the *QESE* scale that were published in books and scholarly/peer-reviewed journals were selected for review. There are several other measures of the self-efficacy beliefs of students learning English as a first language in the literature. These instruments share the basic premises of measuring self-efficacy beliefs, but they differ from one another and from *QESE* in various ways: the domains,

contexts, and tasks being measured, their psychometric properties, and the samples used in validation studies. The following summary of the instruments shows some of those differences. The reader is referred to individual studies for detailed information on each instrument.

4. Results

4.1. Measures of English Self-Efficacy

A search of the literature identified more than 60 measures on the self-efficacy beliefs of learners of English; however, most of them were adapted from the following ten instruments, which is presented in order of publication date.

1. The Reading Self-Efficacy Instrument developed by Shell et al. (1989)

This instrument was developed for college students and consisted of two subscales: a task subscale of 18 reading tasks with varying degrees of difficulty and a skill subscale of nine reading component skills. Students were asked to indicate the probability of successfully performing each of the 18 tasks and their ability to perform each of the nine skills on a scale from zero (no chance) to 100 (completely certain). The internal consistency reliability (Cronbach's α) was reported as .92 for the task subscale and .93 for the component skill subscale. The evidence of criterion-related validity was provided by the correlations between reading achievement and the task self-efficacy ($r = .30$), and reading achievement and the component skill self-efficacy ($r = .53$).

2. The Writing Self-Efficacy Instrument developed by Shell et al. (1989)

This instrument was developed for college students and consisted of two subscales: a task subscale of 16 writing tasks with varying degrees of difficulty and a skill subscale of eight writing component skills. Students were asked to rate their confidence to successfully communicate what they want to say for each of the 16 writing tasks and their ability to perform each of the eight skills on a scale from zero (no chance) to 100 (completely certain). Reliability of this instrument was .92 for the task subscale and .95 for the skill subscale. The evidence of criterion-related validity was provided by the correlations between writing achievement and the task self-efficacy ($r = .17$), and the reading achievement and the component skill self-efficacy ($r = .32$).

3. Self-Efficacy Subscale in Motivated Strategies for Learning Questionnaire (MSLQ) developed by a group of scholars in the University of Michigan and published in Pintrich and DeGroot (1990) study

MSLQ consists of 15 subscales with 56 items, and the 15 subscales can be used separately. Self-efficacy was measured by nine items in the subscale of "Self-Efficacy for Learning and Performance." This instrument was designed to measure undergraduate students' beliefs about how well they can succeed in college. Reliability (Cronbach's α) of the self-efficacy subscale was .89. Factor analysis was conducted to examine the structural aspect of construct validity of this instrument. Since its publication, a group of scholars has examined the psychometric properties of the MSLQ. For example, Pintrich *et al.* (1993); and Sachs *et al.* (2001). MSLQ is by far the most frequently used instrument in the field of educational psychology to measure student beliefs and use of learning strategies based on our literature search, but it is not appropriate in the context of English as a second or foreign language (Wang *et al.*, 2013b).

4. The Reader Self-Perception Scales (RSPS) developed by Henk and Melnick (1995)

Henk and Melnick (1995), developed one general item to prompt children to think about their own reading ability and 32 other items to measure four aspects of self-efficacy in reading: progress, observation comparison, social feedback, and physiological states. The 32 items were written as statements that students were asked to indicate their level of agreement on a 5-point Likert scale. Henk and Melnick (1995), reported the Cronbach's α value for progress as .84, for observation as .82, for social feedback as .81, and for physiological states as .84, indicating satisfactory internal consistency reliability.

5. Writing Self-Efficacy Scale developed by Pajares et al. (2000)

Writing self-efficacy was measured by students' responses about their confidence to earn an A, B, C, D in their language arts class. Although the author claimed that the instrument was developed by following Bandura (1997) guidelines, the items were not provided. As a result, we cannot make a judgement on whether it was a measure of self-confidence or self-efficacy. Pajares *et al.* (2000), reported an internal consistency reliability (Cronbach's α) of .89.

6. Self-Efficacy for Writing Scale developed by Pajares (2007)

Items in this scale were constructed based on the items from the Writing Self-Efficacy Instrument developed by Shell *et al.* (1989), so the scale also consisted of two subscales: basic skills (e.g., spelling, punctuation, and verb tenses) and composition skills (e.g., structuring paragraphs and using topic sentences). The students (Grades 4-11) were asked to make a judgement on their confidence in the basic and composition skills in writing. There was strong evidence for internal consistency reliability as evidenced by Cronbach's alpha ranging from .85 to .92 (Pajares and Valiante, 1997; Pajares *et al.*, 1999; Pajares and Valiante, 1999;2001). The evidence of the structural aspect of construct validity was provided by an exploratory factor analysis, which resulted in a two-factor model. The criterion-related validity was examined by the correlations of its items with some prominent motivation scales (i.e., writing self-concept, self-regulation, writing apprehension, task goal orientation, performance-approach orientation, and performance-avoid orientation) and with students' writing achievement. All the correlations were statistically significant, providing evidence that this instrument had strong criterion-related validity.

7. Self-Efficacy for Listening Comprehension developed by Graham (2007)

This questionnaire tries to measure what students can do in the future by asking students to complete a listening comprehension test before responding to the items in the questionnaire. Students were asked to indicate how confident they were, on a scale from 0 (very unsure) to 100 (completely sure), about tackling a similar task in the

future on four areas which were typical challenges in listening comprehension: (a) understand the gist; (b) understand details; (c) work out the meaning of unknown words; and (d) understand opinions. There is no empirical evidence reported for the reliability or validity of this instrument.

8. Source of English Self-Efficacy Information Scale developed by Chen (2007)

This is a 20-item instrument that measured the four sources of self-efficacy described by Bandura (1986): mastery experience, vicarious experience, social persuasion, and physiological and emotional states. These items were modified versions of the original 40 items developed by Lent *et al.* (1991) in the context of mathematics self-efficacy. The students were asked to rate their level of agreement with the statements from zero (strongly disagree) to 10 (strongly agree). The internal consistency reliability was reported as .92 for mastery experience, .52 for vicarious experience, .76 for social persuasion, and .90 for psycho-physiological state.

9. English Listening Self-Efficacy Measure developed by Chen (2007)

This measure consists of 21 items that ask students to indicate their confidence level on an 11-point rating scale from 0 (not confident at all) to 10 (highly confident), on their capability to perform listening tasks typical in English listening comprehension exams and English courses. The wording of "can" was used in all items to ensure the content aspect of construct validity. For example, one item is "how sure are you that you can listen to and understand the main ideas of a daily conversation between two English speakers." The internal consistency of the response to this measure was .97 Chen (2007).

10. Self-Efficacy in Writing (SEW) developed by Goodman and Cirka (2009)

The instrument consists of 10 items to measure students' self-efficacy in writing, six of which are from the original scale developed by Graham and Harris (1989) and four items developed by Goodman and Cirka (2009). General items were avoided to follow Pajares (1996) suggestion that self-efficacy is task-specific. Goodman and Cirka (2009) adapted the items to reflect better the specific writing tasks in the course they were teaching. The four new items were created by considering the key objectives of the course and what the students were expected to be able to do at the end of the course. For example, a new item is "I find it easy to write a well-organized essay." All items were written in a 7-point Likert scale format. Reliabilities (Cronbach's α) for this instrument were .79 for both the pretest and the posttest. The construct validity of this instrument was examined by a factor analysis, which yielded a unidimensional scale.

4.2. Questionnaire of English Self-Efficacy

The Questionnaire of English Self-Efficacy (QESE) scale includes 32 items developed by Wang (2004). Each item asks students to make judgments about their capabilities to accomplish certain tasks in the context of learning English as a foreign language using listening, speaking, reading, and writing skills. The 32 items were developed with special care through interviews, focus groups, observations, and consultation with experts in educational psychology and second language acquisition (Wang *et al.*, 2013a). Study samples for the studies were collected from China, Korea, Vietnam, United States, and Germany (Huang *et al.*, 2015; Kim *et al.*, 2015; Truong and Wang, 2019; Wang *et al.*, 2012; Wang *et al.*, 2013a). The questionnaire was originally developed with interviews and observations of young Chinese English language learners in the United States, so some modifications were applied to the choice of words in most of the items to reflect the context in each country. According to Bandura (2006), self-efficacy scales must be tailored to the particular domain of interest and the items must be phrased in terms of "can do" instead of "will do" because "will" is a judgment of intention rather than capability. To follow this advice, a conceptual analysis of the language learning process and context was conducted with consultation of English language professionals in each country. The psychometric properties of *QESE* are discussed in the following section.

4.3. Psychometric Properties of QESE

An increasing body of research shows that the *QESE* scale yields valid and reliable inferences for the intended population. The information pertaining to the scale is quite encouraging from measurement perspectives and fills the gap in the literature by providing a reliable and valid instrument to measure English language learners' self-efficacy in different cultures. The following provides summaries of the research findings from studies related to its psychometric properties.

Reliability. Several studies have provided strong evidence for internal consistency as indicated by Cronbach's alpha, ranging from .96 to .99 for all 32 items (Kim *et al.*, 2015; Wang *et al.*, 2012; Wang *et al.*, 2013b; Wang and Kim, 2017). The four areas of listening, speaking, reading, and writing also showed a high degree of internal consistency as indicated by Cronbach's alphas ranging from .88 to .94 (Wang *et al.*, 2014; Wang and Bai, 2017). The scale also showed good test-retest reliability ranging from .81 to .82 (Wang *et al.*, 2013a; Wang and Bai, 2017). Two studies applying the Rasch model (Wang *et al.*, 2013b; Wang *et al.*, 2014) found the scale highly reliable based on the Rasch-based reliability with person reliability and item reliability ranging from .96 to .99.

Evidence of internal structure. Previous studies applying factor analysis resulted in four first-order factors (i.e., listening, speaking, reading, and writing) and a second-order general factor (i.e., general self-efficacy). Wang *et al.* (2013b) found that the two-level four-factor model adequately represented the factorial structure of self-efficacy for both Chinese ($n = 200$) and German ($n = 160$) college student samples. Wang and Bai (2017) also showed that the data from the sample of Chinese secondary school students ($n = 265$) fit the structure of the hypothesized model, providing validity evidence based on the internal structure of the scale.

Two other studies focused on only one single construct (i.e., general self-efficacy) and examined whether the data fit the unidimensional Rasch measurement model. Results from CFA on a sample ($n = 500$) of Chinese sophomore students (Wang *et al.*, 2014) showed that the scale could measure a single latent construct of general

self-efficacy. Results from Rasch analysis on a sample of Korean college students (*n* = 167) also showed that the scale largely satisfied the Rasch model for unidimensionality (Wang *et al.*, 2013b).

Evidence of item-level validity. Two studies examined the item-level psychometric properties of the scale via the Rasch rating scale model (Wang *et al.*, 2013a; Wang *et al.*, 2014). Results from both studies showed that college students in China and Korea reliably distinguished response categories, and the item hierarchy was consistent with the expected item order, supporting evidence for construct validity. The findings suggest that the measure can be used to distinguish individuals with high self-efficacy from those with low self-efficacy. Items related to self-efficacy in listening comprehension were more difficult than those related to self-efficacy in reading, supporting previous findings in the second language learning literature (e.g., (Lund, 1991).

Measurement invariance. Wang *et al.* (2013a), examined measurement invariance by multigroup confirmatory factory analysis, which showed configural, metric, and residual equivalence for college students in China and Germany. The findings suggest that the score interpretation can be generalized across countries, supporting evidence for measurement invariance. German students showed higher self-efficacy than Chinese students did. The results are consistent with the previous findings of cross-cultural studies that academic or general self-efficacy tends to be reported higher among students from individualistic cultures than those from collective cultures (Eaton and Dembo, 1997; Scholz *et al.*, 2002; Schwarzer *et al.*, 1997). Findings from Author (2013a) extended these findings to the domain of learning English as a foreign language.

Criterion-related validity. Results of the structural equation modeling showed that for both Chinese and German students (Wang *et al.*, 2013a), there were significant positive relationships between self-efficacy, use of SRL strategies, and English language test scores, supporting and extending previous research on academic self-efficacy (Pajares and Valiante, 2002; Pape and Wang, 2003; Zimmerman and Martinez-Pons, 1990). Results of path analysis on a sample of Korean college students showed that there were significantly positive relationships between self-efficacy, SRL strategies and English proficiency (Wang *et al.*, 2013b).

Other studies applying Pearson correlation coefficient showed consistent findings, providing evidence for criterion-related validity (Huang *et al.*, 2015; Wang *et al.*, 2012; Wang *et al.*, 2014). More recently, Wang and Bai (2017) found a statistically significant positive correlation between Chinese secondary school students' self-efficacy scores and their English proficiency test scores. A statistically significant positive correlation between student self-efficacy and English proficiency final exam administered six months after the administration of *QESE* provided evidence for predictive validity. These findings support previous studies indicating that self-efficacy is predictive of students' academic achievement (Li and Wang, 2010; Pajares and Valiante, 1997; Pajares and Graham, 1999; Schunk, 1994; Shih and Alexander, 2000).

Empirically identified self-efficacy profiles. Two studies applying latent profile analysis demonstrated that *QESE* could be used to empirically identify underlying latent profiles of ELLs' self-efficacy beliefs. With data from undergraduate students in Kim *et al.* (2015) found that there were three distinct profiles of ELLs' self-efficacy, representing low, medium, and high self-efficacy. The identified self-efficacy profiles were examined in relation to student demographic characteristics, students' use of SRL strategies, and language interpretation strategies and English proficiency. A great number of students in the medium and high self-efficacy profiles were female students and represented those who spent more years studying English. Students with high self-efficacy profile reported more frequent use of SRL strategies and language interpretation strategies, and showed high levels of proficiency in English compared to those with low and medium self-efficacy profiles. The results were consistent with previous findings in the language acquisition as well as self-efficacy literature (e.g., (Pajares, 2009; Pape and Wang, 2003; Wang *et al.*, 2014; Wang and Bai, 2017), which provides evidence for the validity of the *QESE*.

Another study conducted with high school students in China (Wang and Kim, 2017) also found three distinct self-efficacy profiles of ELLs. The findings regarding the relationship between self-efficacy profiles and gender, students' use of SRL strategies, and English proficiency were also consistent with that of the previous study with college students in Korea (Kim *et al.*, 2015). Consistent findings from these studies provide additional evidence for cross-cultural validity.

5. Discussion

This paper highlights our current understanding and measurement of English language learner's self-efficacy beliefs. The studies reported in this paper add to the research literature on ELLs' self-efficacy beliefs by reviewing empirical studies related to ELLs' self-efficacy beliefs as well as by exploring the validity and reliability evidence of *QESE*. A synthesis of empirical studies regarding ELLs' self-efficacy contributes to the self-efficacy theory by expanding the field to English language acquisition. The examination of the psychometric properties of the instruments to measure English language learners' self-efficacy provides useful resources and information for practitioners and researchers in future research. English language classroom teachers will also benefit from this paper through a deep understanding of English language learner's self-efficacy and factors associated with the self-efficacy beliefs, which will guide their curriculum design and pedagogy to enhance the student's learning process and outcomes.

6. Recommended Practices

Traditional English language teaching approaches focus on grammar, reading, vocabulary and writing. Consequently, most past research on ELLs has limited itself to those content areas and traditional indicators of academic language acquisition. However, there is now a growing body of empirical research on ELLs' self-efficacy

and how it relates to other non-cognitive and cognitive skills. Empirical research consistently demonstrates the positive association between ELL's self-efficacy and their English proficiency. Therefore, teachers' understanding of how ELL's self-efficacy affects learning is fundamental to meeting the needs of diverse learners in English language classrooms. The second language curriculum should include measures of self-efficacy as an essential component of assessment. The assessment of self-efficacy can be incorporated into each lesson as a way of individualizing curriculum implementation and providing appropriate instructional interventions to address the specific needs of ELLs.

It is critical that the assessment instruments teachers use are empirically valid. Various instruments to measure self-efficacy beliefs for English language learning have been developed (e.g., (Graham, 2007; Pajares, 2007; Shell *et al.*, 1989; Wang, 2004). Most of these instruments employed Bandura (1986) social cognitive theory, and the items were developed were consistent with Bandura (1986) guidelines. The most common subject areas are reading and writing, although a few were developed to measure self-efficacy in speaking. Only the *QESE* has a subscale to measure ELLs' self-efficacy in speaking. Findings from the studies reported in this paper imply that the *QESE* is a promising instrument to measure ELLs' self-efficacy belief in learning English, providing valid and reliable inferences for the intended population. It is also important that assessment measures are sensitive to differential levels of self-efficacy and diverse populations. Accumulated evidence indicates that the *QESE* can be applied to different cultural groups and contexts, and thus can be used to understand how ELL's self-efficacy is perceived cross-culturally.

It is recommended for teachers to develop high levels of self-efficacy beliefs in the classroom rather than teaching students a set of language learning strategies. Considering cultural variations between collectivistic cultures and individualistic cultures, it is particularly important to teach Asian students to have correct understanding of self-efficacy with accurate judgement of their own capabilities. As such, the *QESE* can be an effective tool for identifying students' self-efficacy beliefs in learning English as a second language. Knowing the level of students' self-efficacy can help teachers plan and individualize instruction for students with varying levels of self-efficacy and English proficiency. Targeted, meaningful feedback that is related to each student's self-efficacy level is better suited to supporting students in learning a second language.

7. Limitations and Future Research

Although the studies on the *QESE* reported in this paper represent data from various groups of students from multiple countries, several limitations must be noted. The major limitation was a small, non-representative sample. There was an uneven distribution of gender. The studies with college students were limited in that the participants majored in a single field of study such as medicine, engineering, business, or education-related fields at top-ranked universities. This may have influenced some of the research findings by precluding students from other degree programs and less prestigious universities. Future studies should aim to include larger samples of students with varying levels of abilities. Studies in the future may consider extending the target population to African American, European, Hispanic cultures, or other Asian cultures.

In addition, the studies only included limited data of student characteristics such as gender, years of study, and age. Certain student information (e.g., students' prior language knowledge and learning experience, motivation, socioeconomic status, and parental involvement) were not available to researchers. Their absence somewhat limited the conclusions drawn from the results. Given the heterogeneity of ELL students, it will be important in future studies to explore students' characteristics including their prior language experiences and parental involvement in relationship to measured items. There is also a need for studies that examine a range of teacher characteristics (e.g., educational level/training, years of teaching experience, and cultural background) and their associations with how teachers use the information from the *QESE* in their instruction as well as its impact on students' learning. This would help researchers and educators understand the roles of teachers in supporting ELL students' second language learning; understanding this relationship could help reduce any potential adverse impacts on students' learning.

The studies used different instruments to measure student English proficiency, most of which were teacher-made or locally developed. This may present difficulties in the generalization of findings and comparisons across studies. Future studies should consider using the same standardized English language proficiency test such as the TOEFL, formerly Test of English as a Foreign Language. This would strengthen conclusions about reliability and validity of the *QESE*. Future studies using longitudinal and experimental approaches may also provide stronger evidence of validity.

To conclude, this paper provides a systematic review of existing instruments to measure English language learner's self-efficacy beliefs, especially the *QESE*. The internal consistency of responses to *QESE* ranged from .88 to .94. The test-retest reliability ranged from .81 to .82. The person reliability was .96 whereas the item reliability was .99. All these statistics suggest that responses to the *QESE* are very consistent and reliable. The validity of responses to the *QESE* was also established through multiple measures. The evidence for the structural aspect of the construct validity was provided through confirmatory factor analysis, and the evidence for the generalizability aspect of the construct validity was provided through cross-cultural comparisons with invariance tests. Results of the cross-cultural comparisons were consistent with previous studies (e.g., (Scholz *et al.*, 2002). The external aspect of the construct validity was examined through latent profile analysis and Pearson correlations, the results of which echoed related studies (e.g., (Lund, 1991; Pajares and Valiante, 2002; Wang *et al.*, 2013b).

In sum, a growing body of literature has begun to document encouraging evidence of ELL students' self-efficacy belief measures and the utility of the *QESE* in particular. The findings from the studies reported in this paper show that self-efficacy beliefs are essential to improving ELL students' learning in English. This paper contributes to

the growing interest in non-cognitive skills in the field of second language learning by reviewing existing studies and instruments employed to understand English language learner's self-efficacy beliefs.

References

Amogne, D. (2008). *An investigation of the correlation among efficacy sources, students' self-efficacy and performance in reading and writing skills: Bahir Dar University in focus (Unpublished master's thesis).* Addis Ababa University: Addis Ababa, Ethiopia.

Anastasiou, D. and Michail, D. (2013). Exploring discourse between self-efficacy and writing performance among low-literate adult students. *Learning Disabilities: A Contemporary Journal,* 11(1): 53-87.

Bai, B., Chao, G. C. and Wang, C. (2018). The relationship between social support, self-efficacy, and English language learning achievement in Hong Kong. *TESOL Quarterly,* 53(1): 208-21.

Baldwin, R. S., Johnson, D. and Peer, G. G. (1980). *BJP reading attitude survey.* Educational Development Corporation: Tulsa, OK.

Bandura, A. (1986). *Social foundations of thought and action.* NJ: Prentice-Hall: Englewood Cliffs.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* NY: Freeman: New York.

Bandura, A. (2006). *Guide for constructing self-efficacy scales. In F. Pajares and T. Urdan (Eds.), Self-efficacy beliefs of adolescents.* Information Age Publishing: Greenwich, CT. 307-37.

Boekaerts, M. and Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review,* 18: 199-210. Available: http://dx.doi.org/10.1007/s10648-006-9013-4

Bong, M. (2006). *Asking the right question: how confident are you that you could successfully perform these tasks? In F. Pajares, and T. Urdan (Eds.), Self-efficacy beliefs of adolescents.* Information Age Publishing: Greenwich, CT. 287-305.

Chen, H. Y. (2007). *The relationship between EFL learners' self-efficacy beliefs and English performance (Unpublished doctoral dissertation).* Florida State University: Tallahassee, Florida.

Corkette, J., Hatt, B. and Benevides, T. (2011). Students and teacher self-efficacy and the connection to reading and writing. *Canadian Journal of Education,* 34(1): 65-98.

Eaton, M. and Dembo, M. H. (1997). Differences in the motivational beliefs of Asian American and non-Asian students. *Journal of Educational Psychology,* 89(3): 433-40.

Goodman, S. B. and Cirka, C. C. (2009). Efficacy and anxiety: An examination of writing attitudes in a first-year seminar. *Journal on Excellence in College Teaching,* 20(3): 5-28.

Graham, S. (2007). Learner strategies and self-efficacy: Making the connection. *Language Learning Journal,* 35(1): 81-93.

Graham, S. and Harris, K. R. (1989). Components analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology,* 81(3): 353-61.

Henk, W. A. and Melnick, S. A. (1995). The reader self-perception scale (RSPS): A new tool for measuring howchildren feel about themselves as readers. *The Reading Teacher,* 48(3): 470-82.

Huang, P., Wang, C. and Xie, L. (2015). *Self-efficacy beliefs and self-regulated learning strategies of chinese students learning english as a second language and as a foreign language. In C. Wang, W. Ma, and C. Martin (Eds.), chinese education from the perspectives of american educators: Lessons learned from study-abroad experiences.* Information Age Publishing: Charlotte, NC. 179-97.

Ivars, A. J., Catalayud, D. P. and Forés, M. R. (2014). Self-efficacy and language proficiency in interpreter trainees. *The Interpreter and Translator Trainer,* 8(2): 167-82.

Kim, D. H., Wang, C., Ahn, H. S. and Bong, M. (2015). English language learners' self-efficacy profiles and relationship with self-regulated learning strategies. *Learning and Individual Differences,* 38(1): 136-42.

Lent, R. W., Lopez, F. G. and Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology,* 38(4): 424-30.

Li, Y. and Wang, C. (2010). An empirical study of reading self-efficacy and the use of reading strategies in the Chinese EFL context. *Asian EFL Journal,* 12(2): 144-62.

Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *The Modern Language Journal,* 75(2): 196-204.

Messik, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement Issues and Practice,* 14(4): 5-8. Available: https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Mills, N. (2004). *Self-efficacy of college intermediate French students: Relation to motivation, achievement, and proficiency (Doctoral dissertation).* Retrieved from ProQuest Dissertations and Theses. (Accession or Order No. 3123349.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research,* 66(4): 543-78.

Pajares, F. (2007). Empirical properties of a scale to assess writing self-efficacy in school contexts. *Measurement and Evaluation in Counseling and Development,* 39(4): 239-49.

Pajares, F. (2009). *Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk and B. J. Zimmerman (Eds.), Motivation and self-regulated learning: Theory, research, and applications.* Routledge: New York, NY. 111-39.

Pajares, F. and Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *The Journal of Educational Research,* 90(6): 353-60.

Pajares, F. and Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology,* 24(4): 390-405.

Pajares, F. and Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology,* 24(2): 124-39.

Pajares, F. and Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology,* 26(3): 366-81.

Pajares, F. and Valiante, G. (2002). Students' self-efficacy in their self-regulated learning strategies: A developmental perspective. *Psychologia,* 24: 211-21. Available: http://dx.doi.org/10.2117/psysoc.2002.211

Pajares, F., Miller, M. D. and Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *Journal of Educational Psychology,* 91(1): 50-61.

Pajares, F., Britner, S. L. and Valiante, G. (2000). Relationship between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology,* 25(4): 406-22.

Pajares, F., Hartley, J. and Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development,* 33(4): 214-21.

Pape, S. J. and Wang, C. (2003). Middle school children's strategic behavior: Classification and relation to academic achievement and mathematical problem solving. *Instructional Science,* 31(6): 419-49. Available: http://dx.doi.org/10.1023/A:1025710707285

Pintrich, P. R. and DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology,* 82(1): 33-40.

Pintrich, P. R., Smith, D. F., Garcia, T. T. and McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement,* 53(1): 801-13.

Prat-Sala, M. and Redford, P. (2010). Writing essays: Does self- efficacy matter? The relationship between self-efficacy in reading and in writing and undergraduate students' performance in essay writing. *Educational Psychology,* 32(1): 9-20. Available: http://dx.doi.org/10.1080/01443410.2011.621411

Rahimi, A. and Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-ROYAL,* 3(1): 14-28.

Sachs, J., Law, Y. K., Chan, C. K. K. and Rao, N. (2001). A nonparametric item analysis of the motivated strategiesfor learning questionnaire–Chinese version. *Psychologia,* 44(3): 197-208.

Sarshar, N. and Oroji, M. R. (2016). The relationship between intermediate EFL students' oral reproduction and their willingness to communicate and self-efficacy. *International Journal of Emerging Trends in Science and Technology,* 3(6): 4149-56.

Scholz, U., Doña, B. G., Sud, S. and Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment,* 18(3): 242-51.

Schunk, D. H. (1994). *Self-regulation of self-efficacy and attributions in academic settings. In D. H. Schunk and B. J. Zimmerman (Eds.), Self-regulation of learning and performance: Issues and educational applications.* US: Lawrence Erlbaum Associates, Inc.: Hillsdale, NJ, . 75-99.

Schwarzer, R., Born, A., Iwawaki, S. and Lee, Y. M. (1997). The assessment of optimistic self-beliefs: Comparison of the Chinese, Indonesian, Japanese, and Korean versions of the general self-efficacy scale. *Psychologia: An International Journal of Psychology in the Orient,* 40(1): 1-13.

Shang, H. (2010). Reading strategy use, self-efficacy and EFL reading comprehension. *The Asian EFL Journal Quarterly,* 12(2): 18-42.

Shell, D. F., Murphy, C. C. and Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology,* 81(1): 91-100.

Shell, D. F., Colvin, C. and Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology,* 87(3): 386-98.

Shih, S. and Alexander, J. M. (2000). Interacting effects of goal setting and self- or other-referenced feedback on children's development of self-efficacy and cognitive skill within the Taiwanese classroom. *Journal of Educational Psychology,* 92(3): 536-43.

Smith, J. K., Smith, L. F., Gilmore, A. and Jameson, M. (2012). Students' self-perception of reading ability, enjoyment of reading and reading achievement. *System,* 22: 202-06. Available: http://doi:10.1016/j.lindif.2011.04.010

Solheim, O. J. (2011). The impact of reading self-efficacy and task value on reading comprehension scores in different items formats. *Reading Psychology,* 32(1): 1-27.

Tremblay, P. F. and Gardner, R. C. (1995). Expanding the motivation construct in language learning. *The Modern Language Journal,* 79(4): 505-18.

Truong, T. N. N. and Wang, C. (2019). Understanding Vietnamese college students' self-efficacy beliefs in learning English as a foreign language. *System,* 84(1): 123-32.

Vogt, C. (2003). An account of women's progress in engineering: A social cognitive perspective. *Journal of Women and Minorities in Science and Engineering,* 9(3/4): 217-38.

Wang, C. (2004). *Self-regulated learning strategies and self-efficacy beliefs of children learning English as a second language. (Doctoral dissertation).* Ohio State University. http://drc.ohiolink.edu/handle/2374.OX/4603

Wang, C. and Pape, S. J. (2007). A probe into three Chinese boys' self-efficacy beliefs learning English as a second language. *Journal of Research in Childhood Education,* 21(4): 364-77.

Wang, C. and Kim, D. H. (2017). *English-language learners in Chinese high schools: Self-efficacy profiles. In M. C. W. Yip (Ed.), Cognition, metacognition and academic performance: An East Asian perspective.* NY: Routledge: New York. 27-39.

Wang, C. and Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly,* 51(4): 931-47.

Wang, C., Schwab, G., Fenn, P. and Chang, M. (2013a). Self-efficacy and self-regulated learning strategies for English language learners: Comparison between Chinese and German college students. *Journal of Educational and Developmental Psychology,* 3(1): 173-91.

Wang, C., Kim, D. H., Bong, M. and Ahn, H. S. (2013b). Examining measurement properties of an English Self-Efficacy Scale for English language learners in Korea. *International Journal of Educational Research,* 59(1): 24-34.

Wang, C., Kim, D. H., Bai, R. and Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System,* 44(1): 24-33.

Wang, C., Hu, J., Zhang, G., Chang, Y. and Xu, Y. (2012). Chinese college students' self-regulated learning strategies and self-efficacy beliefs in learning English as a foreign language. *Journal of Research in Education,* 22(2): 103-35.

Wong, M. S. (2005). Language learning strategies and language self-efficacy: Investigating the relationship in Malaysia. *Regional Language Centre Journal,* 36(3): 245-69.

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System,* 39(4): 510-22. Available: http://doi:10.1016/j.system.2011.10.017

Yang, N. D. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System,* 27(4): 515-35.

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication,* 35(1): 32-57.

Zhang, Y. and Guo, H. (2012). A study of English writing and domain-specific motivation and self-efficacy of Chines EFL learners. *Pan-Pacific Association of Applied Linguistics,* 16(2): 101-21. Available: http://web.a.ebscohost.com

Zimmerman, B. J. and Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology,* 82(1): 51-59.

Zimmerman, B. J. and Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal,* 31(4): 845-62.

Zimmerman, B. J. and Cleary, T. J. (2006). *Adolescents' development of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In F. Pajares and T. Urdan (Eds.), Self-efficacy beliefs of adolescents.* Information Age Publishing: Greenwich, CT. 45-69.

Zuo, H. and Wang, C. (2016). Understanding sources of self-efficacy of Chinese students learning English in an American institution. *Multicultural Learning and Teaching,* 11(1): 83-112.