

Constructing and validating the self-efficacy scale for English language learners' textbooks through Rasch measurement model

Hamedi, Seyyedeh Mina ✉

Ferdowsi University of Mashhad, Iran (sm3_hamedi@yahoo.com)

Pishghadam, Reza

Ferdowsi University of Mashhad, Iran (rpishghadam@yahoo.com)

Ghazanfari, Mohammad

Ferdowsi University of Mashhad, Iran (m.ghazanfari23@gmail.com)



ISSN: 2243-7754
Online ISSN: 2243-7762

OPEN ACCESS

Received: 25 February 2014

Revised: 16 March 2014

Accepted: 18 March 2014

Available Online: 20 March 2014

DOI: 10.5861/ijrsl.2014.732

Abstract

The present study examined the construct validity of Self-Efficacy Scale for English Language Learners' Textbooks (SES-ELLT) using Rasch measurement model. To this end, a sample of 150 language learners from various language institutes participated in the study for the investigation of psychometric properties of SES-ELLT, particularly the reliability and construct validity of the scale. The participants were asked to rate the importance of 36 self-efficacy items for the English language textbooks. Through the Rasch analysis scale reliability, item fit, unidimensionality, and category functioning were investigated. Consequently, the findings demonstrated a fairly high reliability both for the person separation (0.81), and the item separation (0.87). Moreover, three items were detected as misfitting and were discarded, accordingly. In all, the scale exhibited a good unidimensionality and the items sufficiently covered the range of person abilities; however, a 4-category rating scale seemed to be a better alternative as the third and fourth thresholds did not have the required advance between step calibrations in a 5-category rating scale.

Keywords: self-efficacy; textbooks; Rasch measurement model; reliability; construct validity

Constructing and validating the self-efficacy scale for English language learners' textbooks through Rasch measurement model

1. Introduction

Perceived self- efficacy is defined as people's beliefs about their capabilities for learning and performing actions at designated levels (Bandura, 1997). In other words, self- efficacy is defined as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). In fact, perceived efficacy plays a key role in human functioning, because it affects behavior not only directly, but by its influence on other constructs such as goals and aspirations, affective proclivities, outcome expectations, perception of barriers, and opportunities in the social environment (Bandura, 1995, 1997). Self-efficacy has been shown to play a critical role in the individuals' motivation, achievement, and self-regulation (Bandura, 1997; Multon, Brown, & Lent, 1991). It is widely observed that a high sense of self-efficacy for a given task would result in setting higher goals, taking new strategies, and being less afraid of obstacles; whereas, a low sense of it would make the learner avoid the task thoroughly or give it up completely (Bandura, 1997; Zimmerman, 1995).

In the context of Second Language Acquisition (SLA), since textbooks are the major components of most language programs (Richards, 2001) and are considered as an all-important part of language instruction as they provide the essential input through various activities, readings, and explanations (Tomlinson, 2001), one glaring gap that needs to be filled seems to be the evaluation of English Language Teaching (ELT) textbooks regarding their power to enhance the learners' sense of self-efficacy. To the researchers' best knowledge, no study has yet been conducted to construct a self-efficacy scale for the ELT textbooks. Accordingly, the primary goal of the study is to construct and validate a self-efficacy scale for ELT textbooks in order to examine how far English as a Foreign Language (EFL) textbooks can promote the learners' sense of self-efficacy in acquiring a foreign language.

2. Literature Review

2.1 The concept of self- efficacy

The concept of self-efficacy was first proposed by Bandura in 1977 in order to provide a unified theory of behavior change (Gallagher, 2012). Depending on Schunk's (1991, as cited in Koehler, 2007) view, efficacy is the backbone of behavioral change. Meta-analyses across various domains confirm the crucial role of perceived self-efficacy in human self-development, adaptation, and change (Multon et al., 1991). As a matter of fact, efficacy can be promoted through four major channels including mastery experiences, vicarious experiences, social persuasion, and physiological and emotional status (Bandura, 1995). The most effective way of developing a strong sense of self-efficacy is through mastery experiences. Mastery experiences are our direct experiences that are highly considered to be the most informative source of efficacy (Woolfolk, Winne, & Perry, 2003).

The second channel of creating and reinforcing self-efficacy, is through vicarious experiences when people notice a similar individual succeeded by maintaining effort in the face of adversities which may encourage them that they, too, have the essential capabilities to fulfill the same task in a relatively comparable situation (Bandura, 1995). "Usually the more closely the student identifies with the model, the greater the impact on self-efficacy" (Woolfolk et al., 2003, p. 372). The third mechanism by which self-efficacy beliefs can be strengthened is social persuasion. It's commonly referred to as a "peptalk" or "specific performance feedback" (Woolfolk et al., 2003). People are likely to generate effort in the face of setbacks when they are assured verbally that they own the

required capabilities to manage the given activities (Bandura, 1995). The potency of the verbal persuasion may rely on how contented and trusting you are with the persuader (Oettingen, 1997). The more credible, trustworthy, and resourceful one is, the better outcome would be produced (Bandura, 1986). Finally, the last method of generating self-efficacy is on the basis of somatic or emotional cues. This is the least helpful method of developing self-efficacy beliefs, but can still be beneficial in many critical situations (Gallagher, 2012). In fact, people would partly depend on their emotional states in estimating their own capabilities. In other words, they would attribute their stress and apprehensiveness to their susceptibility to poor performance (Bandura, 1995).

Within the context of education, numerous researchers have found that self-efficacy has a stronger impact on academic performance, than other motivational beliefs (Lent, Brown, & Larkin, 1987; Pintrich & De Groot, 1990; Pintrich & Schunk, 2002). Self-efficacy has a great impact on the students' task persistence, task interest, the choices they make and the goals they set (Hosseini Fatemi, Pishghadam, & Vahidnia, 2013). Several studies have explored the role of self-efficacy in mathematics (e.g., Chen, 2003; WanJaafar & MohdAyub, 2010). Within the frameworks of mathematic self-efficacy investigations, WanJaafar and MohdAyub (2010) conducted a study on a group of Malaysian university respondents. As a result, descriptive findings demonstrated that there is a positive correlation between mathematics performance and mathematics self-efficacy.

Recently, there is a growing body of empirical research on the impact of self-efficacy beliefs on second language acquisition. Ching (2002) revealed that students with high self-efficacy beliefs set themselves higher goals, showed greater effort, and were highly resilient and ascribed their failure to their insufficient effort or lack of knowledge. Besides, there has been an increasing attention towards the relationship between language learning strategies, proficiency, and self-efficacy beliefs. Magogwe and Oliver (2007) within the analysis of these three constructs found that, there is a dynamic relationship between language learning strategies, proficiency, and self-efficacy beliefs. They maintained that self-efficacy beliefs relating to language learning "mediate the effect of other influences, such as aptitude or previous achievement, on subsequent performance" (p. 341). More importantly, in the context of SLA, textbooks might be seen as a *mediating object* between the teacher and the learner (Littlejohn & Windeatt, 1989, as cited in Johnson, 1989) as they reveal the objectives of the language program, the kind of syllabus applied, the skills being taught, and the adopted methodologies; however, no study seems to have been conducted to assess the degree to which ELT textbooks can affect the learners' sense of self-efficacy.

2.2 *The role of textbooks in language learning*

Textbooks present new activities and approaches in classrooms regulate learning goals, content, and management of language learning; hence, they are considered as an all-important part of language instruction. Hutchinson and Torres (1994, p. 315) maintained that "*The textbook is an almost universal element of ELT teaching which shapes the teaching and learning process by providing the required structure.*" John (2001) stated that a majority of teachers took advantage of textbooks as their main curriculum guide and source of lessons. Sheldon (1988) described textbooks as the heavily utilized tool by teachers and identified three principal reasons for this: a) developing their own classroom materials is undoubtedly difficult and an arduous process for teachers; b) teachers have restricted time in which developing new materials might not be operationalized; and c) external pressure which confines many teachers in introducing their own developed materials. Despite the aforementioned values, there are some potential negative effects of using textbooks, since they may provide learners with inauthentic language, distorted content, and may ignore students' needs or deskilled teachers (Richards, 2001). Brophy's (1982) research discovered that the school teachers in America did not depend exclusively on their textbooks. Instead, they adapted particular curriculum to the needs of their students in order to make them capable of acquiring life skills.

2.3 *Life skills*

In brief, life skills are the indispensable skills for surviving, living with others, and thriving in a complex

society (Danish & Nellen, 1997). Life skills take in group of skills and abilities which promote individual's far efficient resistance and also in tackling life situations and confliotions (Smith et al., 2004). Essentially, UNICEF proposed an act entitled "teaching life skills" in August 1993, enacted in many countries. It includes many components, each being an individual skill like ability of problem solving, sympathizing with others, creative thinking, decision making, confronting with excitements, confronting with stress, critical thinking, having effective communication, self-awareness, and adaptive inner-person relations (Bashardoost & Zandi, 2010). These skills enable the individual to act adaptively and right in connection with environment and provide self-esteem. Since life skills include the ten mentioned skills all revolving around the concept of ability, boosting self-efficacy seems to be an integral part of any life skill program.

As a result, Pishghadam (2011) proposed life syllabus as a new syllabus which prioritizes life issues in the language syllabus to meet the requirements of both teachers and learners, and to make language learning and teaching more purposeful. With the establishment of life syllabus (Pishghadam & Zabihi, 2012) learners of English can take advantage of language classes by internalizing skills essential for life, therefore "ELT for life" (Pishghadam & Naji, 2012). In sum, although students are the main end-users of ELT textbooks, there seems to be a scarcity of research on the evaluation of EFL textbooks from the students' perspective regarding their power of affecting the learners' sense of self-efficacy; thus, this study aims at constructing and validating a self-efficacy scale for the ELT textbooks based on the learners' view of the currently employed textbooks in Iran.

2.4 *Self-efficacy measurement*

There is no *one measure fits all* of perceived self-efficacy since it usually has restricted explanatory and predictive value as most of the items in an all-purpose scale may not be pertinent to the domain of functioning (Bandura, 2006). Besides, to serve all purposes, items in such a measure usually reflect general terms detached from the situational demands and circumstances; therefore, this leads to much obscurity about what exactly is being measured or the task level and situational demands that must be taken into account (Bandura, 2006). Henceforth, scales of perceived self-efficacy must be in tune with the particular domain of functioning that is the object of interest. During recent decades, self-efficacy has been receiving growing attention in educational research (Dinther, Dochy, & Segers, 2011). Numerous academic self-efficacy scales have been constructed (Chemers, Garcia, & Hu, 2001; Elias & Loomis, 2000; Owen & Froman, 1988; Wood & Locke, 1987); however, most research on self-efficacy in academic settings has been focused on the self-efficacy scale developed by Bandura (Bandura, 2001; Pintrich & De Groot, 1990; Valentine, DuBois, & Cooper, 2004; Zimmerman, 1995). This scale is designed for various academic subjects (e.g., arithmetic) as well as activities associated with general performance at school.

In order to measure the students' self-efficacy and achievement in the English course Templin, Guile, and Okuma (2001) created a reliable and valid self-efficacy scale. In all, in spite of all the efforts to develop self-efficacy scales in various domains, there seems to be a dearth of research to design a reliable and valid self-efficacy scale for the EFL textbooks; henceforth, the major aim of this study is to investigate the construct validity of the SES-ELLT by resorting to Bandura's major sources of self-efficacy and taking the purpose of ELT activities into account. To this end, Rasch analysis is employed in order to obtain an objective measure that does not depend on either the characteristics of the measuring instrument or the skills of individuals (Di Nisio, 2010). Therefore, the present study aims at answering the following question:

- Does SES- ELLT enjoy the acceptable index of reliability and validity?

3. Methodology

3.1 *Participants*

A sample of 150 EFL learners comprising 55 males (36.7%) and 95 females (63.3 %) participated in this

study for the scale validation, with no expectation of incentives. They were all Persian learners of English who were randomly selected from various language institutes of Mashhad, Iran. Indeed, they were high intermediate and advanced students who had a longstanding familiarity with their language textbooks. In all, the ages of the participants ranged from 15 to 64 years. The overall mean age was 26.63 years, with a standard deviation of 9.28.

3.2 Instruments

SES-ELLT

SES-ELLT on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) was constructed and validated through Rasch rating scale model (RSM) (Andrich, 1978). In truth, the items were mainly developed based on previous research studies on self-efficacy, notably the guidelines provided for constructing self-efficacy scales by Bandura (2006) as he is the one who introduced self-efficacy as the main component of social cognitive theory and seeks to measure it by designing task specific scales.

ELT textbooks

In an attempt to construct (SES-ELLT) a total of 150 high intermediate and advanced learners from a range of different institutes of Mashhad were asked to voice their opinions on their current English textbooks, that is, *Top Notch 3B*, *Summit series*, *Interchange 3*, *American English File 4*, and *Total English (advanced)*. They are all communicative English series focusing on both fluency and accuracy. Indeed, they are multi-skills textbooks designed for international communication using natural language. To summarize, they provide a syllabus which integrates themes, grammar, functions, vocabulary, and pronunciation.

Rasch analysis

The Rasch model utilizes a mathematical model in the assessment and testing of construct validity of scales (Tennant & Conaghan, 2007). Its application has extended not only to the dichotomous data, but also to the polytomous data such as Likert-type rating and ordinal data (Bond & Fox, 2007). It has been employed in designing most of the needs-based quality of life instruments (Tennant, McKenna, & Hagell, 2004) to examine the psychometric testing of scale unidimensionality, differential item functioning ([DIF] item bias), as well as constructing dimensions of ordinal scale data, and conversion of ordinal data into interval-level data (Williams, Onsmann, & Brown, 2012). Indeed, it has separate estimates for item difficulty and persons abilities; however, on an interval scale (Pishghadam, Baghaei, & Shayesteh, 2012). Henceforth, the present study resorts to Rasch model to substantiate the psychometric properties of self-efficacy scale for the English language learners' textbooks.

3.3 Procedure

The participants took part in the study from various private language institutes of Mashhad. They were supposed to assess their current language textbooks (e.g., Top Notch, American English Files, Interchange Series, Total English, etc.) with regard to the degree to which their textbook activities can potentially boost their sense of self-efficacy. Indeed, the respondents had to indicate their extent of agreement on a newly designed 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). They were assured of the confidential nature of the study and that they could disengage from the study prior to submitting the questionnaire.

3.4 Data analysis

Firstly, RSM (Bond & Fox, 2007), WINSTEPS software (version 3.74.0) was employed to examine item fit and unidimensionality as the two major aspects of the scale. This is typically measured using the infit mean-square residual fit statistic (MNSQ) as the indication of discrepancies. According to Bond and Fox (2007), the range of acceptable fit should be between 0.70 to 1.30; hence, the misfitting items have to be discarded

following an iterative process since the deletion of one item modifies where the other items sit within the model (e.g., the removal of a difficult item makes the easier items seem somewhat less easy in the model) (Williams et al., 2012). In addition, the unidimensionality of the scale was probed by conducting a Rasch residual-based PCA, to make sure that no other unexpected latent dimensions were contaminating the results. Thereafter, the distributions of item difficulty and person ability measures were further inspected to ascertain whether the items can properly cover the range of person abilities represented by this sample. Finally, the rating scale structure effectiveness is investigated to examine the required advances between the step calibrations.

4. Results

4.1 Reliability measures

As a matter of fact, Rasch reliability statistics were fairly high both for the person separation (.81) (Table 1), and the item separation (.87) (Table 2).

Table 1

Summary of 150 Measured Person

Real RMSE	.21	True SD	.43	Separation	2.07	Person Reliability	.81
Model RMSE	.19	True SD	.44	Separation	2.30	Person Reliability	.84

Table 2

Summary of 36 Measured Item

	Total score	Count	Measure	Model error	Infit MNSQ		
MAX	564	150	.65	.10	1.60		
MIN	428	150	.48	.09	.65		
Real RMSE	.10	True SD	.25	Separation	2.59	Item Reliability	.87
Model RMSE	.09	True SD	.25	Separation	2.70	Item Reliability	.88

4.2 Infit Indices

Afterwards, the Infit statistics, as shown in Table3, identified three items (Items 26, 29, and 32) as misfitting (MNSQ fit values >1.3), with Item 26, “In a listening practice, although I understand almost each word, the main problem is that I do not have the ability to get a whole picture of them in my mind,” (having infit MNSQ 1.31), Item 29, “It makes English listenings a stressful task for me,” (having infit MNSQ 1.59), and item 32, “It makes reading English texts a stressful task for me,” (having infit MNSQ 1.41). It is vital to note that all the three misfitting items were negatively worded items (reverse coding items).

Table 3

Item Estimate and Fit Statistics

Entry number	Total count	Measure	Model SE	Infit MNSQ
26	150	.65	.09	1.31
29	150	.38	.09	1.59
32	150	.15	.09	1.41

Moreover, a Principal Components Analysis (PCA) of Rasch residuals (Linacre, 1998) was conducted to confirm the unidimensionality of the scale. According to Linacre (2006a), a small (usually less than 2.0) eigenvalue of the first contrast reveals that the residuals are random noise, while a greater (usually more than 2.0) eigenvalue means that there is the likelihood of a “second dimension” besides the Rasch dimension. Given the fact that, at a minimum, 50% of the variation should be explained by this measure and the first contrasts explaining more than 5% of the variance or having eigenvalues greater than 2 substantiate the existence of

multiple dimensions and attributes (Wuang, Lin, & Su, 2009); henceforth, as more than 50% of the variance was explained by this measure, and the eigenvalue of the first contrast was less than 2, it revealed that the items in the scale measure a single latent trait. Thereafter, from both theoretical and empirical perspectives, SES-ELLT can be considered as a unidimensional scale.

Additionally, the effectiveness of the distributions of item difficulty and person ability measures were further evaluated as persons and items are placed on the same interval scale, which allows direct comparisons of person and item measures in the Rasch model. To this end, item difficulty and person ability measures were plotted along the latent trait (Figure 1) which means the items with higher difficulty level were placed at the top, and the items with lower difficulty were placed at the bottom; thus, item 26 “In a listening practice, although I understand almost each word, the main problem is that I do not have the ability to get a whole picture of them in my mind” was the most difficult, and item 1 “The current textbook helps me think about my language capabilities” was the least difficult one (Table 2).

As Figure 1 demonstrates, all individuals are clustered towards the center of the scale and the items spread all over the scale. It reveals that there are adequate items in the region of the scale where the persons lie and the construct of self-efficacy is measured along a wide range of abilities.

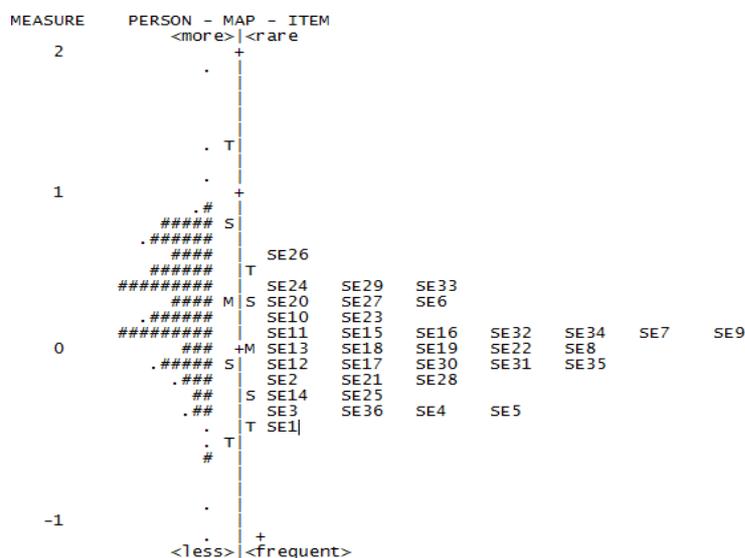


Figure 1. Items-persons map

4.3 Category Functioning

Table 4

Rating Scale Statistics

Category label	Observed count	Observed average	Infit MNSQ	Outfit MNSQ	Andrich threshold	Category measure
1	193	-.08	1.08	1.10	None	-2.84
2	825	0.2	1.00	1.00	-1.53	-1.26
3	1530	.20	.92	.90	-.49	-.17
4	2323	.48	.97	.97	-.07	1.18
5	529	.78	1.02	1.01	2.09	3.27

Ultimately, rating scale structure effectiveness was inspected. In accordance with Linacre and Wright (1998), the step calibrations (the intersection points of adjacent probability curves) of the rating scale must raise monotonically to ascertain that higher measures on the items represent higher traits under measurement. As expected, the observed average should increase with category values. Linacre (2002) notifies that the essential

extent of advance in step difficulties decreases as the number of categories increases; accordingly, he propounds the advances of at least 1.4 logits between step calibrations for a 3-category scale, and for a 5-category rating scale, advances of at least 1.0 logits between step calibrations are required. As shown in Table 4, although thresholds advanced with the categories, they were close together for the third and fourth thresholds (< 1.0 logits apart), suggesting that the participants did not reliably distinguish between these categories; that is, respondents who endorsed “*neitherdisagree nor agree*” (Category 3) on average had the same ability estimate as did respondents who endorsed “*agree*” (Category 4). Therefore, it is recommended to collapse categories 3 and 4 as they had the same ability measure (12334) to improve the rating scale diagnostics.

5. Discussion

The SES-ELLT was developed to investigate the psychometric properties of the scale and was validated through Rasch analysis. In fact, the scale was validated through Item Response Theory (IRT) method (see Appendix). To this end, following Rasch analysis, scale reliability, unidimensionality, the distribution of person-item map, and the appropriateness of the response categories were examined. To sum up, the scale demonstrated a good internal consistency as indicated by the Rasch reliability statistics. The results suggested that the scale items exhibited a good fit to the Rasch model except for three items: item 26 “In a listening practice, although I understand almost each word, the main problem is that I do not have the ability to get a whole picture of them in my mind”, item 29 “It makes English listening a stressful task for me”, and item 32 “It makes reading English texts a stressful task for me”. It is worth noticing that although most of the scale items contributed to a single underlying construct, the problematic and misfitting items proved to be reversing coded ones. This might have occurred by the respondents’ misinterpretation of these items and hence resulted in the item misfit. Possibly, these reverse-scored items might have confused the respondents and this led to the response style bias; thus, defiling the validity of the items.

Besides, the word “them” in the poorly written item 26 seems to be quite ambiguous and confusing as it is not clear whether it is referring to “each word” or the “listening practice”. Afterwards, it is apparent that the learners did not ascribe their stress to the textbook properties since item 29 and 32 were detected as the misfitting ones. Thereafter, these three items were discarded, respectively. Besides, the analysis of dimensionality suggested that the scale measures satisfy the model as more than 50% of the variance was explained by this measure, and the eigenvalue of the first contrast was less than 2. Additionally, examination of the person-item map distribution, which facilitates direct comparisons between person abilities and item difficulty, revealed that the items in the scale were well targeted to the learners’ abilities as the scale items reasonably covered a wide continuum of self-efficacy. Afterwards, the functioning of response categories was inspected and the results proved that the 5-category structure did not function well for the SES-ELLT. It seemed that the third category is subsumed under the fourth category and a better rating structure (a 4-category structure) is worth further investigation. To sum up, the overall results indicated that depending on the Rasch analysis, the SES-ELLT produced a best fit model of 33 items; thus, the scale appeared to enjoy the acceptable index of reliability and validity.

Indeed, previously Bandura (1977, 1986) developed scales to measure the perceived academic efficacy as a part of microanalytic procedure to assess its level across different contexts, but the present study is noteworthy on the premise that no specific scale is yet constructed exclusively to assess ELT textbooks regarding their degree of promoting learners’ self-efficacy. Moreover, in the light of life syllabus proposed by Pishghadam (2011), which gives priority to life issues and language for life purposes rather than merely focusing on the language in class for the sake of preparing the learners to face life challenges, promoting the learners’ whole person development, fostering the learners’ motivation, and confidence, the researchers hope to provide a better context for EFL teachers and learners to fulfill their maximum potential with the aid of enhancing their sense of self-efficacy through ELT textbooks. Furthermore, since the selection of an appropriate ELT textbooks can make a big contribution to the teaching and learning process and is determined by a considerable professional, financial, or even political investment, the findings of the current study will have significant implications for

language teachers, supervisors, L2 learners, lesson planners, material developers, syllabus designers, decision makers, and the language scholars to make valid judgments about the effect of the ELT materials on promoting the L2 learners' sense of self-efficacy and to discriminate between all the available textbooks on the market, accordingly. However, the findings of this study must be considered cautiously with several limitations in mind. Firstly, the sample size was limited due to the strict policies of the language institutes on evaluating their ELT adopted textbooks, at hand. Consequently, the sample of the study could have been larger to come up with more valid generalizations from the results. Secondly, the researchers did not have the opportunity of taking gender effect on self-efficacy promotion into consideration due to the sample size limitation. Thirdly, only a few generations of widely used EFL textbooks were described due to the paucity of language institutes in employing other distinguished textbooks in their syllabi. Finally, the participants were all from language institutes (informal setting) of Mashhad, rather than public schools (formal setting) due to the centralized policies in Iran's public schools; thus, further investigation could have taken place in schools and universities, as well.

6. References:

- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573.
<http://dx.doi.org/10.1007/BF02293814>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215. <http://dx.doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1995). *Self-efficacy in changing societies*. New York: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511527692>
- Bandura, A. (1997). *Self-Efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (2001). *Guide for constructing self-efficacy scales (revised)*. Retrieved May 9, 2005, from <http://www.emory.edu/EDUCATION/mfp/Bandura/>
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age Publishing.
- Bashardoost, F., & Zandi, Z. (2010). Creativity comparison between students who studied life skills courses and those who didn't. *Procedia Social and Behavioral Sciences*, 5, 1390-1395.
<http://dx.doi.org/10.1016/j.sbspro.2010.07.294>
- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.
- Brophy, J. E. (1982). How teachers influence what is taught and learned in classrooms. *The Elementary School Journal*, 83(1), 1-12. <http://dx.doi.org/10.1086/461287>
- Chemers, M. M., Garcia, B. F., & Hu, L.-T. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology*, 93(1), 55-64.
<http://dx.doi.org/10.1037/0022-0663.93.1.55>
- Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences*, 14, 79- 92.
- Ching, L. C. (2002). Strategy and self-regulation instruction as contributors to improving students' cognitive model in an ESL programme. *English For Specific Purposes*, 13, 261-289.
[http://dx.doi.org/10.1016/S0889-4906\(01\)00008-4](http://dx.doi.org/10.1016/S0889-4906(01)00008-4)
- Danish, S. J., & Nellen, V. C. (1997). New roles for sport psychologists: Teaching life skills through sport to at risk youth. *Quest*, 49, 100-113. <http://dx.doi.org/10.1080/00336297.1997.10484226>
- Di Nisio, R. (2010). Measure school learning through Rasch Analysis: The interpretation of results. *Social and Behavioral Sciences*, 9, 373-377.
- Dinther, M. V., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review*, 6, 95-108. <http://dx.doi.org/10.1016/j.edurev.2010.10.003>
- Elias, S. M., & Loomis, R. J. (2000). Using an academic self-efficacy scale to address university major

- persistence. *Journal of College Student Development*, 41, 450-454.
- Gallagher, M. W. (2012). Self-efficacy. In V. H. Ramachandran (2nd ed.), *Encyclopedia of human behavior* (pp. 314-320). Elsevier.
- Hosseini Fatemi, A., Pishghadam, R., & Vahidnia, F. (2013). The role of goal setting theory on Iranian EFL learners' motivation and self-efficacy. *International Journal of Research Studies in Language Learning*, 3(2), 69-84.
- Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal*, 48(4), 315-328 .
<http://dx.doi.org/10.1093/elt/48.4.315>
- John, St. M. (2001). *The status of high science programmes and curricular decision-making*. Inverness, CA: Inverness Research Associates.
- Johnson, K. E. (1989). *The second language curriculum*. New York: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9781139524520>
- Koehler, A. (2007). *Raising awareness of self-efficacy through self-regulated learning strategies for reading in a secondary ESL classroom*. Minnesota: Hamline University.
- Lent, R. W., Brown, S. D., & Larkin, K. C. (1987). Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. *Journal of Counseling Psychology*, 34, 293- 298. <http://dx.doi.org/10.1037/0022-0167.34.3.293>
- Linacre, J. M. (1998). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement*, 2(3), 266-283.
- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3, 85-106.
- Linacre, J. M. (2006a). *A user's guide to WINSTEPS/MINISTEP: Rasch-model computer programs*. Chicago, IL: Winsteps.com.
- Linacre, J. M., & Wright, B. D. (1998). *A user's guide to Bigsteps/Winsteps*. Chicago, IL: Mesa Press.
- Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System*, 35, 338-352.
<http://dx.doi.org/10.1016/j.system.2007.01.003>
- 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, 30-38.
<http://dx.doi.org/10.1037/0022-0167.38.1.30>
- Oettingen, G. (1997). Cross- cultural perspective on self-efficacy. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 149-176). New York: Cambridge University Press.
- Owen, S. V., & Froman, R. D. (1988). *Development of a college academic self-efficacy scale*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.
- Pintrich, P. R., & De Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
<http://dx.doi.org/10.1037/0022-0663.82.1.33>
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and Applications* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Pishghadam, R. (2011). Introducing applied ELT as a new approach in second/foreign language studies. *Iranian EFL Journal*, 7(2), 8-14.
- Pishghadam, R., Baghaei, P., & Shayesteh, S. (2012). Construction and validation of an English language teacher creativity scale (ELT-CS). *Journal of American Science*, 8(3), 497-508.
- Pishghadam, R., & Naji, E. (2012). Applied ELT as a panacea for linguistic imperialism. *Iranian EFL Journal*, 8(1), 35-52.
- Pishghadam, R., & Zabihi, R. (2012). Life syllabus: A new research agenda in English language teaching. *Perspectives*, 19(1), 23-27.
- Richards, J. C. (2001). *The role of textbooks in a language program*. New York: Cambridge University Press.
- Sheldon, L. E. (1988). Evaluating English language teaching textbooks and materials. *ELT Journal*, 42(4), 237-246. <http://dx.doi.org/10.1093/elt/42.4.237>
-

- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26(3 & 4), 207-231. <http://dx.doi.org/10.1080/00461520.1991.9653133>
- Smith, E. A., Swisher, J. D., Vicary, J. R., Bechtel, L. J., Minner, D., Henry, K. L., & Palmer, R. (2004). Evaluation of life skills training and infused-life skills training in a rural setting: Outcomes at two years. *Journal of Alcohol and Drug Education*, 48(1), 51-70.
- Templin, S. A., Guile, T. C., & Okuma, T. (2001). *Creating a reliable and valid self-efficacy questionnaire and English test to raise learners' L2 achievement via raising self- efficacy*. Paper presented at International Conference Center, Kitakyushu, Japan.
- Tennant, A., & Conaghan, P. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism*, 57(8), 1358-1362. <http://dx.doi.org/10.1002/art.23108>
- Tennant, A., McKenna, S., & Hagell, P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. *Value in Health*, 7, 22-26. <http://dx.doi.org/10.1111/j.1524-4733.2004.7s106.x>
- Tomlinson, B. (2001). Materials development. In R. Carter, & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 66-71). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511667206.010>
- Wan Jaafar, W. M., & MohdAyub, A. F. (2010). Mathematics self-efficacy and meta-cognition among university students. *Procedia Social and Behavioral Sciences*, 8, 519-524. <http://dx.doi.org/10.1016/j.sbspro.2010.12.071>
- Williams, B., Onsmann, A., & Brown, T. (2012). A Rasch and factor analysis of a paramedic graduate attribute scale. *Evaluation & the Health Professions*, 35(2), 148-168. <http://dx.doi.org/10.1177/0163278711407314>
- Woolfolk, A. E., Winne, P. H., & Perry N. E. (2003). *Educational psychology*. Toronto: Pearson education.
- Wuang, Y., Lin, Y., & Su, C. (2009). Rasch analysis of the Bruininks-Oseretsky test of motor proficiency-second edition in intellectual disabilities. *Research in Developmental Disabilities*, 30(6), 1132-1144. <http://dx.doi.org/10.1016/j.ridd.2009.03.003>
- Zimmerman, B. (1995). Self-efficacy and educational development. In Bandura, A. (Ed.), *Self- efficacy in changing society* (pp. 202-231). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511527692.009>

Appendix

Self-Efficacy Scale for the English Language Learners' Textbooks

No.	Items	Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree
1.	The current textbook helps me think about my language capabilities.					
2.	It helps me estimate my language capabilities before starting a task.					
3.	It excites my feelings to start doing the tasks.					
4.	It helps me feel I can do my homework effectively.					
5.	It helps me feel I can solve the problems effectively.					
6.	It helps me think how well I can answer the difficult questions in the class.					
7.	It helps me feel I can do my classwork effectively.					
8.	It helps me think how well I can achieve my academic goals.					
9.	It makes me think how well I am doing, as I am proceeding a task.					

10.	It helps me feel determined how to solve a task before I begin.					
11.	It helps me feel confident that I can understand the most difficult materials offered by the textbook.					
12.	It makes me resort to my past experiences while performing a task.					
13.	It encourages me to work as hard as possible on tasks.					
14.	It helps me feel confident that I can understand the basic concept taught by the book.					
15.	It helps me think how well I can master skills by the course.					
16.	It helps me set some of my classmates as the language learner models while doing pair works or group works.					
17.	It encourages me to do extra work on tasks to improve my knowledge.					
18.	It helps me stick to my aims and accomplish my goals.					
19.	It makes me try to use details in my writings to support my ideas.					
20.	It makes me try to write a well-organized, focused text with an inviting beginning, developed middle, and a meaningful ending.					
21.	It makes me try to write well-constructed sentences.					
22.	It makes me try to use effective words in my writings.					
23.	It makes me try to write effectively to express my thoughts and interact with others.					
24.	It makes me try to use punctuation accurately in my writings.					
25.	It helps me feel confident that I have the required ability for improving my listening skill.					
26.	It helps me feel how well I can find a strategy to answer most of the questions even if the listening tasks are hard and I cannot understand them completely.					
27.	It helps me feel confident that my listening comprehension is improving.					
28.	It helps me feel confident that I have the ability to focus my concentration on the text I am reading.					
29.	It helps me feel confident that I am capable of improving my reading comprehension skill.					
30.	It makes me feel although my world knowledge is good, I have problems in reading comprehension.					
31.	It helps me confident that I can understand difficult passages in the textbooks.					
32.	It can make me confident that I can well participate in a class discussion.					
33.	It helps me feel confident that I can communicate my agreement or disagreement in a discussion.					