Developing and Validating a Self-efficacy Questionnaire for CEFR A2 Level EFL Learners[1]

Son Van Nguyen[2,3] & Anita Habók[4,5], University of Szeged, Szeged, Hungary

Abstract

This study provides evidence of the validity and reliability of a self-efficacy questionnaire in English as a Foreign Language (EFL) in a Vietnamese sample of university students. A total of 656 non-English-major students completed the questionnaire in an online format. SPSS version 24, SPSS AMOS, and SmartPLS 3 were employed to analyze the data. A confirmatory factor analysis was carried out to assess the hypothesized structure model and several analyses such as composite reliability, Cronbach's alpha α, and rho_A were conducted to evaluate reliability. The results showed that the questionnaire had a high reliability and adequate validity. All the values were satisfactory, and the model was confirmed. Therefore, the adapted questionnaire in this study can be applied to measure self-efficacy in EFL contexts as it gives valuable feedback to teachers and students to improve the quality of language teaching and learning.

Resumen

Este estudio proporciona evidencia de la validez y confiabilidad de un cuestionario de autoeficacia en inglés como lengua extranjera (ELE) en una muestra vietnamita de estudiantes universitarios. Un total de 656 estudiantes que no estudiaban la carrera de inglés completaron el cuestionario en un formato en línea. Se emplearon SPSS versión 24, SPSS AMOS y SmartPLS 3 para analizar los datos. Se llevó a cabo un análisis factorial confirmatorio para evaluar el modelo de estructura hipotetizado y se realizaron varios análisis como la confiabilidad compuesta, el alfa de Cronbach y rho_A para evaluar la confiabilidad. Los resultados mostraron que el cuestionario tenía una alta fiabilidad y una validez adecuada. Todos los valores fueron satisfactorios y se confirmó el modelo. Por lo tanto, el cuestionario adaptado en este estudio se puede aplicar para medir la autoeficacia en los contextos de inglés como lengua extranjera y, como resultado, brinda una valiosa retroalimentación a los profesores y estudiantes para contribuir a mejorar la calidad de la enseñanza y el aprendizaje de idiomas.

Introduction

In the 1970s, several scholars recommended that studies should be done on the non-linguistic outcomes of second/foreign language learning (Wang et al., 2013). This resulted in studies focusing on issues such as motivation (e.g., Bai & Wang, 2021; Clément et al., 1985; Dörnyei, 2001), language learning strategies (e.g., Anam & Stracke, 2016; Habók & Magyar, 2018a; Montaño-González & Cancino, 2020; O'Malley & Chamot, 1990; Oxford, 1990; 2011), and self-regulation (e.g., Bai & Wang, 2021; Habók & Magyar, 2018b; Tseng et al., 2017; Schunk & Zimmerman, 2012; Zimmerman, 2011). Researchers have also examined self-efficacy, which is considered one of the most important factors in motivation (Bandura, 1997; Hoang & Wyatt, 2021). Self-efficacy is believed to greatly influence students' academic performance in myriad ways such as language proficiency, learning motivation, and mathematical achievement (e.g., Habók & Magyar, 2020; Habók et al., 2020; Hoang & Wyatt, 2021; Mills et al., 2007; Pajares & Graham, 1999; Shih & Alexander, 2000).

Researchers assert that self-efficacy, like many other constructs in educational psychology, can be applied to the field of second/foreign language education (Tseng et al., 2006). Studies demonstrate a positive correlation between self-efficacy and language skills, including writing (Golparvar & Khafi, 2021; Rayner et al., 2016; Wang & Bai, 2017; Woodrow, 2011), listening (Graham, 2011; Mills et al., 2008), reading (Ghonsooly & Majid, 2010; Wang & Pape, 2007), speaking (Aregu, 2013), and language proficiency in general (Bai et al., 2019; Diseth, 2011; Liem et al., 2020; Wang & Sun, 2020).
However, research on self-efficacy in language learning is under-developed, especially in the context of Vietnam (e.g., Hoang, 2020; Hoang & Wyatt, 2021; Nguyen & Phan, 2020; Phan & Locke, 2016; Truong & Wang, 2019). Although research instruments have been developed to investigate EFL learners' self-efficacy (e.g, Wang, 2004), they are neither validated in detail nor suited for the Vietnamese context of language education (e.g., Truong & Wang, 2019). Therefore, the purpose of this study was to examine the psychometric properties of an instrument measuring self-efficacy among EFL Vietnamese learners.

Theoretical Background

Self-efficacy is defined as "belief in one's capacity to organize and execute the course of action required to manage prospective situations" (Bandura, 1997, p. 2). This involves learners' judgments about their own ability to utilize their skills to accomplish a task. There are four distinct features of self-efficacy, including: (1) perceived capacity to perform an activity; (2) specific to the task, domain, or context; (3) a criterion of high performance; and (4) an afore-thought process (measured before any task engagements) (Zimmerman & Cleary, 2006). Specifically, the second feature of self-efficacy infers that, measures of self-efficacy vary across particular tasks within a specific domain, thereby being multidimensional in nature. The third feature implies that students are not required to compare their work to other people's, but rate their own capabilities (Wang et al., 2014). These key characteristics facilitated the process of constructing and validating the questionnaire in this study.

Self-efficacy concepts stem from social cognitive theory developed by Bandura (1986), which highlights triadically reciprocal relationships among the personal, behavioral, and environmental factors that determine behaviors. According to this theory, a human being is cognitively able to organize, reflect, and regulate their own behavior. Thus, self-efficacy is an influential factor in human behavior, according to social cognitive theory (Mills et al., 2007; Wang, et al., 2013). Self-efficacy beliefs influence what learners do, the way they

[1] This is a refereed article. Received: 8 March, 2021. Accepted: 21 July, 2021. Published: 1 January, 2022.
[2] Doctoral School of Education
[3] nvson@tlu.edu.vn, nguyen.son.van@edu.u-szeged.hu, 0000-0002-2614-7831, Correspondent.
[4] Institute of Education, University of Szeged; MTA-SZTE Digital Learning Technologies Research Group
[5] habok@edpsy.u-szeged.hu, 0000-0003-0904-8206

think, self-motivate, feel, and behave (Bandura, 1997). Such beliefs "powerfully influence the level of accomplishment that one ultimately achieves" (Pajares, 2008, p. 113).

Highly self-efficacious students tend to be more responsible for their learning process, more willing to accept challenging tasks, more persistent in dealing with obstacles, more flexible in using learning strategies, and more accurate in self-evaluating their academic performance (Kim et al., 2015; Mills et al., 2008; Schunk, 2012; Zimmerman & Kitsantas, 2005). These students also exhibit lower levels of anxiety (Schunk, 1983). By contrast, students with low levels of self-efficacy are likely to avoid complicated tasks and not take up challenges (Mills el al., 2007; Schunk, 1990). Therefore, self-efficacy contributes to predicting learners' academic success (Bandura, 1997; Shih & Alexander, 2000), so developing psychometrically sound scales to measure self-efficacy is increasingly valuable.

Self-efficacy has not only been researched in the context of English as a first language, but also in the context of EFL (Wang & Sun, 2020). Studies indicate that self-efficacy is a contributing factor in predicting achievement in listening (e.g., Graham, 2011; Rahimi & Abedini, 2009), reading (e.g., Naseri & Zaferanieh, 2012; Shang, 2010), speaking (e.g., Asakereh & Dehghannezhad, 2015; Zhang et al., 2020), and writing (e.g., Yavuz Erkan & Iflazogl Saban, 2011; Schunk & Swartz, 1993). However, there is a lack of EFL self-efficacy scales, and scholars have sometimes not assessed self-efficacy appropriately (Bong, 2006). Consequently, a more accurate assessment tool of self-efficacy in learning English as a foreign language is needed (Wang et al., 2013).

Literature Review

As previously noted, the literature on self-efficacy is extensive. However, the studies on self-efficacy in the EFL context are limited (e.g., Kim et al., 2015; Zheng et al., 2017). In addition, instruments that measure EFL self-efficacy are also lacking. This section discusses previous studies on the validation of measures for EFL self-efficacy. The review does two things: (1) identifies the research gap; and (2) indicates the current study's contribution to the literature on EFL self-efficacy. Several criteria were used to include and exclude studies for review:

- Type: Journal articles
- Language: English
- Date: from 2010 to present
- Context: EFL
- Databases: the list of journals in SCOPUS
- Keywords: self-efficacy questionnaires, self-efficacy scales for language learners, validate, psychometric properties.

A search using the keywords listed above, resulted in three studies.

Wang et al. (2013) adapted the Questionnaire of English Self-Efficacy (QESE) that Wang (2004) developed to investigate the psychometric properties on a sample of 167 university Korean students. They used the Rasch rating scale model to analyze the data and verify the item hierarchy. The model allowed "item parameters to be sample-free and person parameters to be item-independent" to enhance generalizability beyond the sample (p. 28). The results showed that QESE largely fulfilled unidimensionality and had adequate reliability. However, the items did not cover the entire continuum of the variables observed, hence the QESE required more proof of validity.

In another study, Wang et al. (2014) validated the QESE originally established by Wang (2004) based on internal consistency reliability, construct validity (including convergent validity as well as criterion-related validity), and the Rasch rating scale model. The sample comprised 500 second-year students at a university in the Southeast of China. The findings showed that QESE had acceptable reliability and validity. The item hierarchy was consistent with the item order expected, which facilitated the scale's construct validity.

Wang and Bai (2017) also examined psychometric properties of two instruments to measure EFL self-efficacy and self-regulated learning strategies, one of which was the QESE. Messick's (1995) framework of validity was used to validate the QESE using a sample of 265 students from a Chinese secondary school. The authors conducted reliability analyses on internal consistency, test-retest, correlation analysis (with English proficiency), and confirmatory factor analysis (CFA). The results indicated that the QESE was a highly reliable research instrument with an adequate level of validity. Students' responses fit the hypothesized model of four factors.

Although the QESE can be used in EFL contexts, more studies are needed to demonstrate its reliability for researchers studying EFL self-efficacy beliefs. In the QESE, there are a total of 32 can-do questions that require students to judge their capabilities to do a specific activity when they learn English measured by four skills—listening, reading, writing, and speaking. The scale was developed following Bandura's (1997) theory and covered these four skills of the language learning process. Contextually, the QESE has been validated only through limited samples in China and Korea; thus, it is necessary to provide more evidence of the validity of the QESE using more samples in different countries. Truong and Wang (2019) explored a sample of Vietnamese students' EFL self-efficacy using the QESE but did not thoroughly describe the analytical procedures nor provide much evidence of reliability except for internal consistency. Therefore, a more detailed validation study in the context of Vietnam is needed. Methodologically, the above-reviewed studies utilized different approaches to validate the QESE, but all found the questionnaire to be psychometrically sound. Additional ways to investigate reliability and validity, such as composite reliability, rho_A reliability, and CFA with more cut-off values, would reinforce the QESE's value.

In conclusion, the present study will fill research gaps by carrying out validation procedures of an adapted version of the QESE on a Vietnamese sample and by applying additional methods to investigate the questionnaire's properties. The study will add to the literature on EFL self-efficacy and provide a potential

tool for educators and researchers to gain a greater understanding of non-linguistic factors in the language learning processes.

Methodology and Data Collection

Participants

A total of 656 undergraduates participated in the study, all second-year students at a large public university in Vietnam. The percentage of males was 57.3% (n = 376) and females was 42.7% (n = 280). The mean age was 19 years. The students' majors were civil engineering (CE, n = 62; 9.4%), economics (ECO, n = 197; 30%), electrical engineering (EE, n = 68; 10.4%), environmental studies (n = 44; 6.7%), information technology (n = 135; 20.6%), mechanical engineering (n = 109; 16.6%), and water resources engineering (n = 41; 6.3%). Their English proficiency was reported to be elementary, although most of them had already studied English for about 10 years. Among the participants, 292 responses were used for exploratory factor analysis (EFA), whereas the other 656 responses were utilized for confirmatory factor analysis.

At the time of data collection, the participants were studying an English course provided by EFL instructors at the university. The textbook used was Prepare! Student's Book and Online Workbook Level 2 (Kosta & Williams, 2015) in which an emphasis was put on language skills (listening, reading, speaking, and writing) integrated with pronunciation, vocabulary, and grammar into each unit. The 90-hour course aimed to enable the students to achieve English level B1, which is the benchmark set by the university.

Instrument

The questionnaire used to measure EFL self-efficacy in this study was the QESE adapted from Wang (2004) who utilized interviews, verbal protocols, and observations from Chinese English language learners in the U.S. That questionnaire was further developed and validated by Wang et al. (2013), Wang et al. (2014), and Wang and Bai (2017). For the current study, some modifications were made so that the questions in the survey would accurately reflect the study's specific context in Vietnamese higher education where English has not been acknowledged as a second language and it is limited to classroom settings. For instance, question 25 in the QESE, which was on self-efficacy in reading, asked students whether they could understand articles about Chinese culture written in English. The question was adapted to "Can you understand descriptions of events, feelings, and wishes in personal letters?" provided by Common European Framework of Reference (CEFR) descriptions of A2 level proficiency (Council of Europe, 2020). This was because students in this investigation did not have access to English articles about Vietnamese culture, but they had been exposed to personal letters in their textbooks and supplementary materials in A2 level which were the main sources of English language. Although the descriptors in CEFR were general for any language learner, they have been used as a framework of benchmarks in the Vietnamese context and they shed light on the fundamentals of the language courses the students were taking. All the changes based on CEFR A2 level are listed in Table 1. A group of six people with expertise in educational psychology and language education at the university level then held three videoconference discussions to evaluate the item pool. Judging representativeness, double-barelled parts, vagueness, and sensitivity, the experts agreed on which questions in the pool would be substitutes for the original ones in Wang (2014).

The original questions in Wang (2004)	The questions in this study
Can you understand radio programs in English-speaking countries?	9. Can you understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc.?
Can you understand English language TV programs made in Korea?	13. Can you catch the main point in short, clear, simple messages, and announcements?
Can you understand numbers spoken in English?	28. Can you understand the English instructions provided by your language teacher?
Can you do homework alone when they include reading English texts?	2. Can you understand familiar names, words, and very simple sentences, for example on notices and posters or in catalogs?
Can you read English newspapers?	17. Can you find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus, and timetables?
Can you understand English articles about Chinese culture?	25. Can you understand the description of events, feelings, and wishes in personal letters?
Can you understand new reading materials (e.g., news from the Time magazine) selected by your English instructor?	29. Can you understand texts that consist mainly of high frequency every day or job-related language?
Can you describe your university to other people in English?	3. Can you use simple phrases and sentences to describe where you live and people you know?
Can you ask your English instructor questions in English?	14. Can you briefly give reasons and explanations for opinions and plans?
Can you introduce your English instructor to someone else in English?	18. Can you ask and answer simple questions in areas of immediate need or on very familiar topics?
Can you leave a note for another student in English?	12. Can you write short, simple notes, and messages relating to matters in areas of immediate needs?
Can you make sentences with English idiomatic phrases?	23. Can you fill in forms with personal details, for example entering your name, nationality, and address on a hotel registration form?
Can you write diary entries in English?	27. Can you write a very simple personal letter, for example thanking someone for something?
Can you write a two-page essay about your English instructor in English?	31. Can you write clear, detailed text on a wide range of subjects related to your interests?

Table 1: Changes in the new questionnaire compared to the original one

There were a total of 32 questions in the questionnaire asking students to judge their capacity to carry out certain EFL tasks (see Appendix A). The questionnaire was designed with a 7-point rating Likert scale ranging from 1 (*I cannot do it at all*) to 7 (*I can-do it well*). Four separate areas were measured: (1) self-efficacy for Listening (questions 1, 5, 9, 13, 16, 20, 24, and 28); (2) self-efficacy for Reading (questions 2, 6, 10,

17, 21, 25, 29, and 32); self-efficacy for Speaking (questions 3, 7, 11, 14, 18, 22, 26, and 30); self-efficacy for Writing (questions 4, 8, 12, 15, 19, 23, 27, and 31). Each question was phrased as a *can-do* question to focus on the capability rather than the intention, as happens with a *will-do* question (Bandura, 2006).

After finishing the English version of the survey, a back-translation method was used to translate the questions into Vietnamese, the native language of the participants, to ensure that they could all understand what the survey was about (see more in Behr, 2017; Brislin, 1970; Sousa and Rojjanasrirat, 2011). This process was supported by several ELT experts, including one British professor, two Vietnamese PhDs from Australian universities, one Vietnamese PhD candidate in New Zealand, and four Vietnamese language instructors holding Master's degrees in TESOL who are based in Vietnam.

Afterwards, three undergraduates at the university who had the same language education background as the participants were invited to discuss the version together online on *Zoom*. They were provided with details on the significance and aims of the study. They commented that the questions were all relevant to their language learning process and they did not find anything strange and challenging.

In addition, three other experts in educational psychology including one from Australia and two from Vietnam were consulted to give their opinions and judgments on the questions in terms of readability, understandability, and relevance to the field and the cultural context. The final Vietnamese version of the questionnaire was then produced.

Data collection procedure

An online form of the survey was created because Internet-based questionnaires tend to be more widely distributed, faster, environmentally friendlier, and cheaper than a paper-and-pencil questionnaire (Cohen et al., 2018). The introduction of the online form provided potential participants with a comprehensive description of the study's aims, significance, ethical issues, and methods. Students were informed that participation in the study would not affect them in terms of grades and their personal lives.

The pilot survey was conducted with five students who would not participate in the actual study to complete the questionnaire and give feedback on its clarity. These students and the participants came from the same cohort. They had the same level of English proficiency as the research participants and they were also attending the participants' English course. These students in the pilot stage did not have any difficulty understanding the format and content, so no changes were needed. They took approximately twenty minutes to fill in the survey. The link was sent to students with the support of their language instructors via class groups on social networks such as Zalo which is a Vietnamese network and Facebook. In the first phase of over a four-week period, from the middle of November 2020 to the middle of December 2020, 387 responses were recorded, 23 of which were not usable due to their incompleteness. This gave an overall response rate of 94 percent. The second phase lasted in four days during the third week of May, 2021 and 301 students returned the surveys but 292 of them were used and nine of them were discarded because they were incomplete. Although data were collected in two phases, the participants' characteristics in terms of language learning such as proficiency, language courses, and syllabus are the same.

Data analysis procedure

The data were coded into SPSS version 24, which was the foundation for data entered into SPSS AMOS and SmartPLS 3. The software generated outputs analyzing reliability and validity.

The former was ascertained based on internal consistency reliability (Cronbach's alpha α), composite reliability (CR), and rho_A reliability. CR of a scale is "the ratio of its true score variance divided by its observed score variance" (Peterson & Kim, 2013, p. 194). Rho_A reliability is assessed by "the off-diagonal elements of a latent variable's indicator correlation matrix are reproduced as well as possible in a least squares sense" (Dijkstra & Henseler, 2015, p. 300).

For validity, CFA was utilized via structural equation modeling to check the model hypothesized regarding EFL self-efficacy. A number of goodness-of-fit indices were used to assess the model fitness, including CFI (comparative fit index), NFI (normed fit index), TLI (Tucker-Lewis index), SRMR (standardized root mean square residual), RMSEA (the root mean square error of approximation), RMS_theta (the root mean square residual covariance), Chi-square χ^2. CFI, NFI, and TLI are incremental comparative fit indices, whereas SRMR, RMSEA, RMS_theta, and normed Chi-square are absolute fit indices. The descriptions of those indices are listed in Table 2.

Indices	Descriptions	Cut-off values
CFI	• investigates whether the proposed model lacks fit over an independence model or not (Kline, 2015; Rigdon, 1996) • is not affected by the model's complexity (Teo et al., 2013)	> = 0.9 (Bentler, 1990; Hooper et al., 2008)
NFI	• works on Chi-square value and then makes a comparison of that value to a meaningful benchmark (Bentler & Bonett, 1980) • is sensitive to sample size (Bentler, 1990)	> = 0.9 > = 0.95 (Hu & Bentler, 1990)
TLI	• is an alternative of NFI • "penalizes the Chi² values by the degrees of freedom (df)"	> = 0.9 (Bentler & Bonett, 1980)
Normed Chi-square	• is calculated by dividing Chi-square by degree of freedom (χ^2/d.f.) (Glynn et al.i, 2011) because Chi-square depends much on sample size (Cangur & Ercan, 2015)	2<χ^2/d.f.<5 ((Byrne, 1991; Tabachnick & Fidell, 2007; Wheaton et al., 1977)
SRMR	• represents the degree of error caused by the estimation of the specified model	< = 0.06 (Teo, 2013)
RMSEA	• "corrects the tendency of the χ^2 to reject models with large same size or number of variables" (Teo et al., 2013, p. 15)	< = 0.05 (a confidence level of 95%)
RMS_theta	• refers to the outer model residuals' root mean squared residual covariance matrix (Henseler et al., 2016; Lohmöller, 1989) • assesses the extent of correlations among outer model residuals	<0.12 (Henseler et al., 2015)

Table 2: Good-of-fitness indices and their values

The construct validity of the questionnaire was assessed by convergent validity and discriminant validity. Convergent validity reflects the relationship between the items and questions in a hypothesized model (Carlson & Herdman, 2010; Cunningham et al., 2001; Habók & Magyar, 2018b), while discriminant validity refers to the degree to which the items/questions in one construct can be differentiated from those in another (Henseler et al., 2015; Hair et al., 2010). Convergent validity is established if the loadings in the same scale/sub-scale are not less than 0.70. Average variance extracted (AVE) should also be more than 0.50 and CRs should be higher than 0.70. However, AVE is sometimes too restricted (Malhotra & Dash, 2011) so, if CR is higher than 0.60, convergent validity can be confirmed. Discriminant validity is established by employing heterotrait-monotrait ratio of correlations (HTMT). This is defined as "the average of the heterotrait-heteromethod correlations (i.e., the correlations of indicators across constructs measuring different phenomena), relative to the average of the monotrait-heteromethod correlations (i.e., the correlations of indicators within the same construct)" (Henseler et al., 2015, p. 121). If HTMT is lower than 0.90, discriminant validity will be ascertained. Moreover, maximum shared squared variance (MSV) is employed to assess discriminant validity. If the values of AVE are greater than 0.5 and higher than those of corresponding MSV, discriminant validity can be affirmed (Almén et al., 2018; Alumran et al., 2014; Hair et al., 2010; Obrad, 2020; Rebelo-Pinto et al., 2014).

In this study, the hypothesized model was based on the categorization of self-efficacy beliefs in EFL into four separate factors including self-efficacy for reading (SER), self-efficacy for listening (SEL), self-efficacy for speaking (SES), and self-efficacy for writing (SEW). CFA would help confirm or reject that hypothesized model.

Results

Exploratory factor analysis

EFA was employed in this study to investigate students' responses because the questionnaire was adapted and developed from an existing one used in other contexts and cultures. Moreover, the items were modified and rewritten so that they were appropriate for the participants' comprehension. The initial statistical tests including KMO and Barlett's test showed that the data were suited for factor analysis ($\chi 2$ = 8647.678; d.f. = 496; $p < 0.001$; KMO = 0.968).

EFA was conducted with the 32 questions (n = 292) on the basis of principal components extraction and varimax rotation with Kaiser normalization because factors were assumed to be independent (Field, 2018). The rotation was converged in eight iterations. The analysis produced four factors with eigenvalues higher than 1 and the solution of these four factors accounted for 70.842% of the total variance. The factor loadings are presented in Table 3.

Sub-scales and questions	Sub-scales			
	SEL	SER	SES	SEW
Self-efficacy for listening				
Q20	*0.803*	0.098	0.242	0.306
Q5	*0.789*	0.193	0.023	0.105
Q28	*0.719*	0.209	0.143	0.150
Q24	*0.711*	0.032	0.302	0.087
Q13	*0.689*	0.003	0.110	0.220
Q16	*0.664*	0.104	0.078	0.240
Q9	*0.645*	0.038	0.079	0.048
Q1	*0.619*	0.232	0.149	0.295
Self-efficacy for reading				
Q29	0.121	*0.800*	0.156	0.234
Q25	0.002	*0.790*	0.301	0.053
Q32	0.028	*0.777*	0.229	0.203
Q10	0.248	*0.770*	0.166	0.114
Q17	0.022	*0.764*	0.078	0.240
Q6	0.261	*0.710*	0.155	0.276
Q2	0.231	*0.694*	0.300	0.170
Self-efficacy for speaking				
Q18	0.179	0.035	*0.826*	0.190
Q14	0.074	0.152	*0.818*	0.279
Q26	0.253	0.082	*0.804*	0.066
Q11	0.306	0.090	*0.785*	0.094
Q7	0.134	0.178	*0.761*	0.176
Q22	0.062	0.210	*0.756*	0.068
Q3	0.094	0.227	*0.705*	0.017
Q30	0.309	0.127	*0.694*	0.067
Self-efficacy for writing				
Q27	0.237	0.073	0.214	*0.828*
Q19	0.199	0.079	0.073	*0.811*
Q12	0.022	0.057	0.257	*0.785*
Q31	0.157	0.111	0.176	*0.762*
Q8	0.121	0.312	0.066	*0.756*
Q15	0.221	0.122	0.053	*0.744*
Q4	0.208	0.311	0.124	*0.710*
Q23	0.301	0.145	0.156	*0.672*

Table 3: Factor loadings in EFA

Question 21 was excluded from the questionnaire because its loading was low in all sub-scales (0.111; 0.316; 0.243; and 0.277). The other questions loaded over 0.6 and their cross-loadings were all lower than 0.32 (Worthington & Whittaker, 2006).

Confirmatory factor analysis

A total of 656 students' responses in a separate sample were used to conduct CFA. The goodness-of-fit indices for the measured questions (CFI = 0.912; NFI = 0.924; TLI = 0.908; χ^2/d.f. = 4.51; SRMR = 0.046;

RMSEA = 0.042; RMS_theta = 0.109) suggest that students' responses fit the structure of the hypothesized model. Question 21 was deleted because its low loading factor caused a deterioration in the reliability of the sub-scale *self-efficacy for reading*. Thus, after the analysis, 31 questions remained. The result of the four-factor model is represented in Figure 1.

In this model, ovals represent latent variables (i.e., self-efficacy for listening, reading, speaking, and for writing), while rectangles represent the extant variables (i.e., question items). Smaller ovals show measurement errors for observed variables. Two-way arrows indicate correlations between variables that are different English skills.

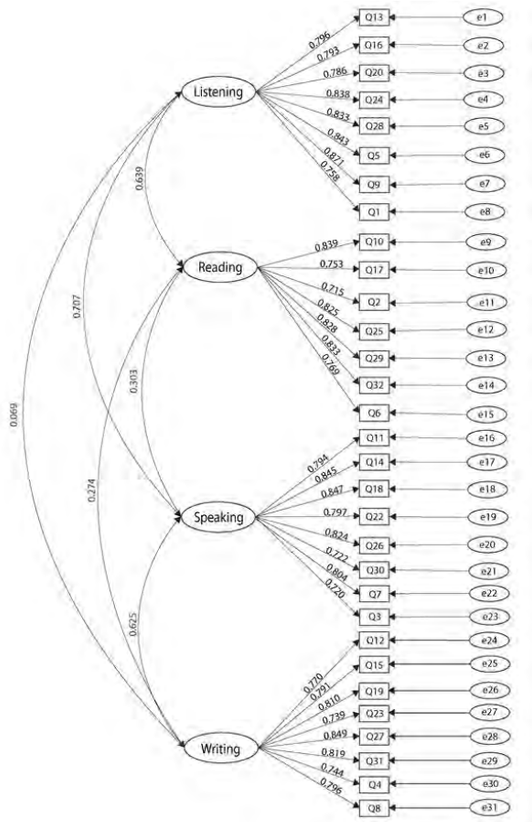

Figure 1: Model for the 31-item self-efficacy scale for EFL learners

Reliability

Analyses of reliability showed that the sub-scales in the survey had excellent levels of reliability, including Cronbach's alpha α, CR, and rho_A reliability (Table 4).

Sub-scales	Cronbach's alpha	CR	Rho_A
Self-efficacy for listening	0.928	0.941	0.929
Self-efficacy for reading	0.903	0.923	0.905
Self-efficacy for speaking	0.916	0.932	0.919
Self-efficacy for writing	0.914	0.930	0.916

Table 4: Values of reliability analyses

Specifically, self-efficacy for listening demonstrated the highest reliability (α = 0.928; CR = 0.941; rho_A = 0.929), while self-efficacy for reading indicated the lowest values of reliability among the four (α = 0.903; CR = 0.923; rho_A = 0.905). The other two fields' reliabilities also attained high levels.

Validity

The evidence regarding content validity and face validity of the questionnaire was collected during its development process. In terms of content validity, the expert panel discussed and reviewed the item pool for decisions to use substitutes or not, comprehensiveness, and clarity. Then, after a version of the questionnaire in the native language was drawn up, some EFL learners used a think-aloud protocol and gave feedback based on its readability. With regard to face validity, some undergraduates helped with clarity of language, as well as the content of each question. The experts in educational psychology double-checked the linguistic quality, the technical aspects, the relevance of each question, and the questions' representativeness. They evaluated whether each question matched the domain of the concept of self-efficacy. All feedback was taken into consideration and they agreed that the questionnaire was a valid measure of self-efficacy beliefs. The students who supported piloting the questionnaire did not face any challenges comprehending and completing the questionnaire.

For convergent validity, the CFA results indicated that all the factor loadings for the questions were higher than 0.70 with $p < 0.001$. Moreover, the values of AVE ranged from 0.625 to 0.665 and CRs were all greater than 0.70 (from 0.923 to 0.941). Thus, convergent validity was established (Table 5).

Sub-scales & questions	Factor loading	AVE	CR	MSV
Self-efficacy for listening (SEL)		0.665	0.941	0.49
Q9	0.871			
Q5	0.843			
Q24	0.836			
Q28	0.833			
Q13	0.796			
Q16	0.793			
Q20	0.786			
Q1	0.758			
Self-efficacy for reading (SER)		0.634	0.923	0.41
Q10	0.839			
Q32	0.833			
Q29	0.828			
Q25	0.825			
Q6	0.769			
Q17	0.753			
Q2	0.715			
Self-efficacy for speaking (SES)		0.632	0.932	0.49
Q18	0.847			
Q14	0.844			
Q26	0.824			
Q7	0.804			
Q22	0.796			
Q11	0.792			
Q30	0.724			
Q3	0.721			
Self-efficacy for writing (SEW)		0.625	0.930	0.39
Q27	0.849			
Q31	0.819			
Q19	0.809			
Q8	0.796			

Q15	0.791
Q12	0.770
Q4	0.744
Q23	0.739

Table 5: CFA of the questionnaire

HTMT ratio was used to assess discriminant validity. The results indicate that all HTMT ratios were less than 0.90, ranging from 0.756 to 0.812 (Table 6). Another index to indicate discriminant validity is MSV. It can be seen in Table 5 that all the MSV values were lower than their AVE counterpart. This means that discriminant validity was confirmed in this study.

Sub-scale	SEL	SER	SES	SEW
SEL				
SER	0.798			
SES	0.781	0.768		
SEW	0.756	0.774	0.812	

Table 6: HTMT ratios

Discussion

The main objective of this study was to develop and validate an adaptation of the QESE. This included questions that measure self-efficacy in English language learning among A2 (CEFR) Vietnamese learners.

The questionnaire was piloted and validated after the first attempt to develop it according to the guidelines for a reliable and valid survey questionnaire (see more in Dörnyei, 2003; Gillham, 2008). The set of cut-off values and fit indices provide ample evidence of both the reliability and the validity of the questionnaire. These results also confirm the structure of the questionnaire, which enabled identification of four different factors, including self-efficacy for listening, reading, speaking, and writing in EFL courses among a sample of students in a university in Vietnam. The proposed model of four factors regarding self-efficacy beliefs among CEFR A2 level EFL learners (i.e., SEL, SER, SES, and SEW) fit with the data after question 21 was excised, which additionally improved reliability and validity of the questionnaire. The final structure of the questionnaire with 31 can-do questions was as follows: SEL (8 questions), SER (7 questions), SES (8 questions), and SEW (8 questions). This structure looked almost identical to previous studies (e.g., Wang et al., 2013; Wang et al., 2014; Wang & Bai, 2017), which meant the questionnaire designed by Wang (2004) proved structurally sound.

Notably, although previous studies employed the questionnaire published by Wang (2004) and made some modifications, the questions' content remained largely unchanged. This study adapted, but neither changed the format of each question nor the number of questions. Some questions were found to be inappropriate in the Vietnamese context, so the content of several questions was localized to ensure that the students were culturally familiar with their English-learning context in Vietnam on the basis of CEFR A2 level. These adaptations enhance the commonsense element that an effective questionnaire requires (Dörnyei & Csizér, 2012). The validation of the questionnaire was also different from studies that utilized Wang's (2004) questionnaire. Several values, including alphas, CR, and rho_A, were applied to double-check whether the instrument was reliable or not. CFA confirmed the structure of the questionnaire and the hypothesized model. More fit indices (e.g., CFI, NFI, TLI, SRMR, RMSEA, and RMS_theta) were used to explore validity compared to other studies, as suggested by Martens (2005). Despite the previously noted changes and differences, the version of the questionnaire used in this study demonstrated high levels of reliability and validity.

With regards to its contribution to EFL literature, this study demonstrates potential for the application of the questionnaire to measure self-efficacy beliefs in the context of EFL. The questionnaire can be particularly useful for EFL teachers in Vietnam, as well as other countries where English is not the native or first language, for understanding students' level of self-efficacy—a vital element in language learning. This is because CEFR A2 level descriptors were employed to redesign questions that mirrored the language learning process of the non-English-major students with A2 level in contexts such as Vietnam. The results can provide useful feedback to teachers, as well as students, on self-efficacy beliefs. This could help improve the quality of non-linguistic language learning in the Vietnamese context, where global integration and internationalization has made EFL increasingly important (Phan, 2021). Besides motivating the students or providing them with learning strategies, teachers are advised to nurture highly self-efficacious language learners in the EFL classrooms (Pajares, 2002). Theoretically, the current study shows that the instrument initially constructed by Wang (2004) and its adapted versions, including the one in this study, are reliable tools for measuring self-efficacy beliefs among EFL learners.

This study has some limitations that deserve attention. First, only online surveys in *Google Forms* were administered to the participants. Therefore, it is difficult to establish whether those who did not complete the survey understood the questionnaire well or not (see more in Cohen et al., 2018). This might explain why many students who attended the course at the university did not participate in this research. Second, only EFA, CFA and reliability analyses were used to validate the questionnaire, and other variables, such as age, gender, language proficiency, academic achievement, socio-economic variables, motivation, and strategy usage, were not taken into account. Such factors should be included in future studies to provide more validity evidence for the questionnaire and ascertain the relationships between self-efficacy beliefs and those factors. Moreover, a qualitative approach was used to investigate content validity, so it is recommended that a quantitative approach such as content validity ratio (e.g., Polit & Beck, 2006; Shrotryia & Dhanda, 2019) be applied in further studies to obtain more objective evidence of content validity. Also, more validation studies on QESE and its versions in different samples are needed to make this a widely used and useful research tool, thereby remedying the scarcity of instruments in the field of EFL self-efficacy.

Conclusion

The adapted QESE in this study was proven to be a reliable and valid instrument to measure EFL self-efficacy beliefs. Based on the guidelines from Bandura (2006), the process started with the systematic review of the literature on self-efficacy, particularly for English language skills. Then, several well-established questionnaires were checked, and the questionnaire constructed by Wang (2004) was found noteworthy. However, not all the questions in that questionnaire were appropriate for Vietnamese students in the sample, so some questions were modified and adapted to the descriptions of A2 English level issued by CEFR. ELT experts and university students facilitated the process of developing the questionnaire. After collecting data from 656 undergraduates at a public university in Hanoi, the data were analyzed using SPSS and SmartPLS3. CFA, fit indices, and reliability values were used to validate the questionnaire. The findings showed that the questionnaire possessed high levels of reliability and validity, demonstrating that this tool can be used by researchers and educators who wish to interrogate self-efficacy beliefs in the field of EFL. Some limitations in the present study were discussed and suggestions made for future research related to validation of the self-efficacy measures.

References

Almén, N., Lundberg, H.., Sundin, Ö., & Jansson, B. (2018). The reliability and factorial validity of the Swedish version of the Recovery Experience Questionnaire. *Nordic Psychology, 70*(4), 324-333. https://doi.org/10.1080/19012276.2018.1443280

Alumran, A., Hou, X.-Y., Sun, J., Yousef, A. A., & Hurst, C. (2014). Assessing the construct validity and reliability of the parental perception on antibiotics (PAPA) scales. *BMC Public Health, 14*(1), 1-9. https://dx.doi.org/10.1186%2F1471-2458-14-73

Anam, S. & Stracke, E. (2016). Language learning strategies of Indonesian primary school students: In relation to self-efficacy beliefs. *System*, *60*, 1–10. http://dx.doi.org/10.1016/j.system.2016.05.001

Aregu, B. B. (2013). Self-efficacy, self-esteem and gender as determinants of performance in speaking tasks. *Journal of Media and Communication Studies*, *5*(6), 64–71. https://doi.org/10.5897/JMCS2013.0366

Asakereh, A., & Dehghannezhad, M. (2015). Student satisfaction with EFL speaking classes: Relating speaking self-efficacy and skills achievement. *Issues in Educational Research, 25*(4), 345–363. http://www.iier.org.au/iier25/asakereh.pdf

Bai, B., Chao, G. C. N., & Wang, C. (2019). The relationship between social support, self efficacy, and English language learning achievement in Hong Kong. *TESOL Quarterly*, 208–221. https://doi.org/10.1002/tesq.439

Bai, B. & Wang, J. (2021). Hong Kong secondary students' self-regulated learning strategy use and English writing: Influences of motivational beliefs. *System*, *96*, 1–14. https://doi.org/10.1016/j.system.2020.102404

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W.H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age.

Behr, D. (2017). Assessing the use of back translation: The shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology, 20*(6), 573–584. https://doi.org/10.1080/13645579.2016.1252188

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–600. https://doi.org/10.1037/0033-2909.88.3.588

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

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Byrne, B. M. (1991). The Maslach Burnout Inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators. *Multivariate Behavioral Research, 26*(4), 583–605. https://doi.org/10.1207/s15327906mbr2604_2

Cangur, S., & Ercan, I. (2015). Comparison of model fit indices used in structural equation modeling under multivariate normality. *Journal of Modern Applied Statistical Methods, 14*(1), 152–167. https://doi.org/10.22237/jmasm/1430453580

Carlson, K. D., & Herdman, A. O. (2010). Understanding the impact of convergent validity on research results. *Organizational Research Methods, 15*(1), 17–32. https://doi.org/10.1177/1094428110392383

Clément, R., Dörnyei, Z., & Noels, K. A. (1994). Motivation, self confidence, and group cohesion in the foreign language classroom. *Language Learning, 44*(3), 417–448. https://doi.org/10.1111/j.1467-1770.1994.tb01113.x

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). Routledge.

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Council of Europe.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163–170. https://doi.org/10.1111/1467-9280.00328

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* Plenum.

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02

Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences, 21*(2), 191–195. https://doi.org/10.1016/j.lindif.2011.01.003

Dörnyei, Z. (2001). *Teaching and researching motivation.* Longman.

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration and processing*. Mahwah, NJ: Erlbaum.

Dörnyei, Z., & Csizér, K. (2012). How to design and analyze surveys in second language acquisition research. In A. Mackey, & S. M. Gass (Eds.), *Research methods in second language acquisition* (pp. 74–94). Wiley.

Field, A. (2018). *Discovering statistics using SPSS* (5th ed.). Sage.

Golparvar, S. E., & Khafi, A. (2021). The role of L2 writing self-efficacy in integrated writing strategy use and performance. *Assessing Writing*, *47*, 1–14. https://doi.org/10.1016/j.asw.2020.100504

Ghonsooly, B., & Majid, E. (2010). Learners' self-efficacy in reading and its relation to foreign language reading anxiety and reading achievement. *Journal of English Language Teaching and Learning,, 53*(217), 45–67. https://profdoc.um.ac.ir/articles/a/1020898.pdf

Gillham, B. (2008). *Developing a questionnaire*. Continuum.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science Motivation Questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 48*(11), 1159–1176. https://doi.org/10.1002/tea.20442

Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic Purposes, 10*(2), 113–117. https://doi.org/10.1016/j.jeap.2011.04.001

Habók, A. & Magyar, A. (2018a). The effect of language learning strategies on proficiency, attitudes and school achievement. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2018.01388

Habók, A., & Magyar, A. (2018b). Validation of a self-regulated foreign language learning strategy questionnaire through multidimensional modelling. *Frontiers in Psychology, 9,* Article e1368. https://doi.org/10.3389/fpsyg.2018.01388

Habók, A., & Magyar, A. (2020). The role of students' approaches in foreign language learning. *Cogent Education*, *7*(1). https://doi.org/10.1080/2331186X.2020.1770921

Habók, A., Magyar, A., B. Németh, M. & Csapó, B. (2020). Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data. *International Journal of Educational Research, 103*, Article 101634. https://doi.org/10.1016/j.ijer.2020.101634

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Prentice Hall..

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2–20. https://www.emerald.com/insight/content/doi/10.1108/IMDS-09-2015-0382/full/html

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hoang, H. T. T.(2020). *Vietnamese EFL teacher self-efficacy in practice.* [Unpublished doctoral dissertaion]. Victoria University of Wellington. http://hdl.handle.net/10063/9284

Hoang, T., & Wyatt, M. (2021). Exploring the self-efficacy beliefs of Vietnamese pre-service teachers of English as a foreign language. *System*, *96*. https://doi.org/10.1016/j.system.2020.102422

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60. https://academic-publishing.org/index.php/ejbrm/article/view/1224/1187

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kim, D.-H., Wang, C., Ahn, H. S, & Bong, M. (2015). English language learners' self-efficacy profiles and relationship with self-regulated learning strategies. *Learning and Individual Differences, 38*, 136–142. https://doi.org/10.1016/j.lindif.2015.01.016

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed). Guilford.

Kosta, J., & Williams, M. (2015). *Prepare! Student's book and online workbook.* Cambridge.

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology, 33*(4), 486–512. https://doi.org/10.1016/j.cedpsych.2007.08.001

Malhotra, N. K., & Dash, S. (2011). *Marketing research: An applied orientation.* Pearson.

Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist, 33*(3), 269–298. https://doi.org/10.1177%2F0011000004272260

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

Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning, 57*(3), 417–442. https://d1wqtxts1xzle7.cloudfront.net/47476481

Mills, N., Pajares, F., & Herron, C. (2008). A reevaluation of the role of anxiety: Self efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276–295. https://doi.org/10.1111/j.1944-9720.2006.tb02266.x

Montaño-González, J., & Cancino, M. (2020). Exploring the relationship between language learning strategies and self-efficacy of Chilean university EFL students. *MEXTESOL Journal*, *44*(2), 1–16. http://www.mextesol.net/journal/public/files/a334ae33e6742db9b3bd687a4540b890.pdf

Naseri, M., & Zaferanieh, E. (2012). The relationship between reading self-efficacy beliefs, reading strategy use and reading comprehension level of Iranian EFL learners. *World Journal of Education, 2*(2), 64–75. https://www.sciedupress.com/journal/index.php/wje/article/view/992/505

Nguyen, H. X., & Phan, N. T. T. (2020). Students' self-efficacy beliefs and TOEIC achievements in the Vietnamese context. *International Journal of Instruction*, *13*(4), 67-86. https://doi.org/10.29333/iji.2020.1345a

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge University Press.

Obrad, C. (2020). Constraints and consequences of online teaching. *Sustainability, 12*(17), 1-23. https://doi.org/10.3390/su12176982

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know.* Newbury House.

Oxford, R. L. (2011). *Teaching and researching language learning strategies.* Longman/Pearson Education.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning*. Theory Into Practice, 41*(2), 116–125. https://doi.org/10.1207/s15430421tip4102_8

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111–139). Routledge.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124–139. https://doi.org/10.1006/ceps.1998.0991

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology, 98*(1), 194–198. https://doi.org/10.1037/a0030767

Phan, A. N. Q. (2021). Under the impacts of globalisation: the rising power of English as a foreign language (EFL) and the corresponding response of EFL policy in Vietnam. *SN Social Sciences*, 1(1). https://doi.org/10.1007/s43545-020-00047-9

Phan, N. T. T., & Locke, T. (2016). Vietnamese teachers' self-efficacy in teaching English as a Foreign Language: Does culture matter? *English Teaching: Practice & Critique, 15*(1), 105-128. https://doi.org/10.1108/ETPC-04-2015-0033

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497. https://onlinelibrary.wiley.com/doi/abs/10.1002/nur.20147

Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-ROYAL, 3*(1), 14–28. https://novitasroyal.org/volume-3-issue-1-april-2009/?wpdmc=volume-3-issue-1#

Rayner, G., Papakonstantinou, T., & Gleadow, R. (2016). Comparing the self-efficacy and writing-related abilities of native and non-native English-speaking students. *Cogent Education*, *3*(1). http://dx.doi.org/10.1080/2331186X.2016.1179164

Rebelo-Pinto, T., Pinto, J. C., Rebelo-Pinto, H., & Paiva, T. (2014). Validation of a three-dimensional model about sleep: Habits, personal factors and environmental factors. *Sleep Science, 7*(4), 197-202. https://doi.org/10.1016/j.slsci.2014.12.002

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 3*(4), 369–379. https://doi.org/10.1080/10705519609540052

Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology, 75*(6), 848–856. https://doi.org/10.1037/0022-0663.75.6.848

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*(1), 71–86. https://doi.org/10.1207/s15326985ep2501_6

Schunk, D. (2012). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol.1: Theories, constructs, and critical issues* (pp. 101–123). American Psychological Association.

Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18*(3), 337–354. https://doi.org/10.1006/ceps.1993.1024

Schunk, D. H., & Zimmerman, B. J. (2012). Self regulation and learning, In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology, Volume 7* (2nd ed.). (pp. 45–68). Wiley.

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

Shih, S.-S., & 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), 536–543. https://doi.org/10.1037/0022-0663.92.3.536

Shrotryia, V. K., & Dhanda, U. (2019). Content validity of assessment instrument for employee engagement. *Sage Open, 9*(1), 1-7. https://doi.org/10.1177/2158244018821751

Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross cultural health care research: A clear and user friendly guideline. *Journal of Evaluation in Clinical Practice, 17*(2), 268–274. https://doi.org/10.1111/j.1365-2753.2010.01434.x

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.

Teo, T. (2013). Examining the psychometric properties of the epistemic beliefs inventory (EBI). *Journal of Psychoeducational Assessment, 31*(1), 72–79. https://doi.org/10.1177/0734282912449436

Teo, T., Tsai, L. T., & Yang, C.-C. (2013). Applying structural equation modeling (SEM) in educational research: An introduction. In M. S. Khine (Ed.), *Application of structural equation modeling in educational research and practice* (pp. 3–22). Sense.

Truong, N. N. T., & Wang, C. (2019). Understanding Vietnamese college students' self-efficacy beliefs in learning English as a foreign language. *System, 84*, 123–132. https://doi.org/10.1016/j.system.2019.06.007

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78–102. https://doi.org/10.1093/applin/ami046

Tseng, W.-T., Liu, H., & Nix, J.-M. L. (2017). Self-regulation in language learning: Scale validation and gender effects. *Perceptual and Motor Skills, 124*(2), 531–548. https://doi.org/10.1177/0031512516684293

Wang, C. (2004). *Self-regulated learning strategies and self-efficacy beliefs of children learning English as a second language* [Unpublished doctoral dissertation] , Ohio State University. http://rave.ohiolink.edu/etdc/view?acc_num=osu1091546670

Wang, C., & Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self efficacy beliefs and self regulated learning strategies. *TESOL Quarterly*, *51*(4), 931–947. https://doi.org/10.1002/tesq.355

Wang, C., Kim, D.-H., Bong, M., & Ahn, H. S. (2013). Examining measurement properties of an English Self-Efficacy scale for English language learners in Korea. *International Journal of Educational Research, 59*, 24–34. https://doi.org/10.1016/j.ijer.2013.02.004

Wang, C., Kim, D.-H., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System, 44*, 24–33. https://doi.org/10.1016/j.system.2014.01.015

Wang, C., & 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), 364–377. https://doi.org/10.1080/02568540709594601

Wang, C., & Sun, T. (2020). Relationship between self-efficacy and language proficiency: A meta-analysis. *System, 95*, 1–11. https://doi.org/10.1016/j.system.2020.102366

Wheaton, B., Muthen, B., Alwin, D. F., & Summer, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology, 8*(1), 84–136. https://doi.org/10.2307/270754

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System, 39*(4), 510–522. https://doi.org/10.1016/j.system.2011.10.017

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838. http://doi.org/10.1177/0011000006288127

Yavuz Erkan, D., & Iflazoglu Saban, A. I. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in the Turkish tertiary-level EFL context. *Asian EFL Journal, 13*(1), 164–192. https://asian-efl-journal.com/PDF/March-2011-dye.pdf

Zhang, X., Ardasheva, Y., & Austin, B. W. (2020). Self-efficacy and English public speaking performance: A mixed method approach. *English for Specific Purposes, 59*, 1–16. https://doi.org/10.1016/j.esp.2020.02.001

Zheng, C., Liang, J.-C., & Tsai, C.-C. (2017). Validating an instrument for EFL learners' sources of self-efficacy, academic self-efficacy and the relation to English proficiency. *Asia-Pacific Education Researcher, 26*(6), 329–340. https://doi.org/10.1007/s40299-017-0352-3

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In: B. J. Zimmerman, & D. H. Schunk (Eds), *Self-regulation of learning and performance (*pp. 49–64). LEA.

Zimmerman, B. J., & Cleary, T. J. (2006). Adolescents' developement of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In F. Parajes, & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 45–69). Information Age.

Zimmerman, B. J., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology, 30*(4), 397–417. https://doi.org/10.1016/j.cedpsych.2005.05.003

Appendix

Dear Students,

My name is …….. I am conducting a study on university students' self-efficacy in language learning. Participating in my research will not cause any harms to you. I would be grateful if you could help me complete the following survey. There are 32 questions in total in the survey. The questions in the survey aim to investigate your judgement of your capacities, so there will be neither right nor wrong answers. It may take you 15-20 minutes to complete it. There is an example for you.

Please read the questions carefully and choose the option of number that best reflects your judgement of language learning capacities.

1: I cannot do it at all.

2: I cannot do it.

3: Maybe I cannot do it.

4: Maybe I can do it.

5: I basically can do it.

6: I can do it.

7: I can do it well.

Example:

Can you introduce yourself in English fluently?	①	②	③	④	❺	⑥	⑦

	Questions							
1.	Can you understand stories told in English?	①	②	③	④	⑤	⑥	⑦
2.	Can you understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues?	①	②	③	④	⑤	⑥	⑦
3.	Can you describe your university to other people in English?	①	②	③	④	⑤	⑥	⑦
4.	Can you compose messages in English on the Internet (Facebook, Twitter, blogs, etc.)?	①	②	③	④	⑤	⑥	⑦
5.	Can you understand TV programs in English?	①	②	③	④	⑤	⑥	⑦
6.	When you read English texts, can you guess the meaning of unknown words?	①	②	③	④	⑤	⑥	⑦
7.	Can you describe the way to the university from the place where you live in English?	①	②	③	④	⑤	⑥	⑦
8.	Can you write a short paragraph assigned by your English instructor?	①	②	③	④	⑤	⑥	⑦
9.	Can you I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc?	①	②	③	④	⑤	⑥	⑦
10.	Can you understand the English news on the Internet?	①	②	③	④	⑤	⑥	⑦
11.	Can you tell a short story in English?	①	②	③	④	⑤	⑥	⑦
12.	Can you write short, simple notes and messages relating to matters in areas of immediate needs?	①	②	③	④	⑤	⑥	⑦
13.	Can you catch the main point in short, clear, simple messages and announcements?	①	②	③	④	⑤	⑥	⑦
14.	Can you briefly give reasons and explanations for opinions and plans?	①	②	③	④	⑤	⑥	⑦
15.	Can you make new sentences with the words just learned?	①	②	③	④	⑤	⑥	⑦
16.	If your English instructor gives you a tape-recorded English dialogue about everyday school matters, can you understand it?	①	②	③	④	⑤	⑥	⑦
17.	Can you find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables?	①	②	③	④	⑤	⑥	⑦
18.	Can you ask and answer simple questions in areas of immediate need or on very familiar topics?	①	②	③	④	⑤	⑥	⑦
19.	Can you write email messages in English?	①	②	③	④	⑤	⑥	⑦
20.	Can you understand English movies without subtitles?	①	②	③	④	⑤	⑥	⑦
21.	Can you find out the meaning of new words by using English–English dictionaries?	①	②	③	④	⑤	⑥	⑦
22.	Can you discuss subjects of general interest with your fellow students in English?	①	②	③	④	⑤	⑥	⑦
23.	Can you fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form?	①	②	③	④	⑤	⑥	⑦
24.	Can you understand English songs?	①	②	③	④	⑤	⑥	⑦
25.	Can you understand English articles about Vietnamese culture?	①	②	③	④	⑤	⑥	⑦
26.	Can you respond to your English instructor's questions in English?	①	②	③	④	⑤	⑥	⑦
27.	Can you write a very simple personal letter, for example thanking someone for something?	①	②	③	④	⑤	⑥	⑦
28.	Can you understand English instructions provided by your language teacher?	①	②	③	④	⑤	⑥	⑦
29.	Can you understand texts that consist mainly of high frequency everyday or job-related language?	①	②	③	④	⑤	⑥	⑦
30.	Can you introduce yourself in English?	①	②	③	④	⑤	⑥	⑦
31.	Can you write clear, detailed text on a wide range of subjects related to my interests?	①	②	③	④	⑤	⑥	⑦
32.	Can you read short English stories?	①	②	③	④	⑤	⑥	⑦