

The Concurrent Validity of Sukhothai Thammathirat Open University's English Proficiency Test with the Computer System under the Common European Framework of Reference for Languages

Sajee Jiraro^{1*} • Supamas Aungsuchoti²

¹ The Office of Registration, Records and Evaluation, Sukhothai Thammathirat Open University, Nonthaburi, Thailand.

² Retired Government Official, Samut Songkhram, Thailand.

*Corresponding author. Email: Sajee.jir@stou.ac.th

Article Info:

Received: 10 May 2022; Revised: 22 Nov 2022; Accepted: 23 Dec 2022; Available Online: 27 Dec 2022

Abstract

This research was conducted to measure the concurrent validity of Sukhothai Thammathirat Open University's English Proficiency Test with the computer system or STOU-EPT (E-Testing) under the Common European Framework of Reference for Languages. The sources of the data were 30 participants who received their results in 2019 from STOU-EPT (E-Testing) and volunteered to sit for the International English Language Testing System (IELTS) and Srinakharinwirot University's Standardised English Test (SWU-SET). The samples for the population were conveniently stratified into three groups according to their level of proficiency in STOU-EPT (E-Testing): A2, B1 and B2, whereby each level had 10 participants. Two electronic surveys were used as the data collection instruments; the first was to collect information to register for IELTS and SWU-SET, and the latter was for data analysis. IBM SPSS statistics software was used as the data analysis instrument to process the participants' results of English proficiency in all three tests. The consistency between the tests was measured according to three aspects, Listening, Structuring, and Reading. The concurrent validity of STOU-EPT (E-Testing) was measured by using Kendall's tau-b formula. It was found that the levels of CEFR of STOU-EPT (E-Testing) were significantly related to IELTS and SWU-SET at the significance level of 0.01 in all aspects between the strength of 0.700 and 0.890. The research was concluded with the confirmation of the concurrent validity of STOU-EPT (E-Testing) with the recommendation for revision in the Listening skill aspect.

Keywords: *Common European Framework of Reference for Languages, computer-based test, concurrent validity test, English Proficiency Test, open university*

1. Introduction

The Common European Framework of Reference for Languages (CEFR) is widely used as a teaching guide for English language instructors in the European Union and other countries as a standardisation of evaluation of linguistic skills; assessing one's proficiency in using the English language for communication and cultural inclusivity. CEFR was found to help support ongoing progress in language learning (Council of Europe, 2021) and was widely used to determine the minimum linguistic criteria for various purposes. Thus, many countries outside of Europe that use the English Language as a second language or foreign language have referred to CEFR in constructing their own's personalised English proficiency standards to ensure their citizen; at the international level, are marketable and able to represent their country, in addition to ensuring consistent language education countrywide.

In China, for example, when Peng et al. (2022) statistically compared the country's English proficiency standards, China's Standards of English Language Ability and CEFR through Rasch analysis, they found that the standard was consistent with CEFR on multiple assessment levels. Similar results can be found in Malaysia's English language curriculum assessment, whereby Mohamad et al. (2020) found that most of the assessment items found that most of the writing and reading assessments in 331 Malaysian secondary schools were tailored to the CEFR framework. Meanwhile, the Higher Education Commission's Announcement B.E. 2559 outlined a policy requiring local educational institutes to provide English proficiency tests with standards equivalent to CEFR for their students (Thailand's Education Commission, 2016), applicable to all levels of education including at the university level.

Likewise, the policy also applies to Sukhothai Thammathirat Open University (STOU); an open university in Thailand studying in the distance education system. It developed a test to measure the ability to use English at the university accordant to CEFR, referred to as the STOU's English Proficiency Test (STOU-EPT); used for evaluating the English proficiencies of the applicants who want to study for bachelor's degrees at the university since 1999. To ease the process of the assessment, the university's Office of Registration, Records and Evaluation started developing an online course and computer-based testing to replace the paper-based test in 2010. The electronic version of the test is called STOU-EPT (E-Testing), first fully implemented in 2017 (Jiraro and Angsuchoti, 2021). The conversion successfully eliminated the issues of limitation or wastage in terms of the storage, venue, and materials associated with STOU-EPT.

To ensure that STOU-EPT (E-Testing) has a structure consistent with CEFR, it was developed by instructors with experience in creating tests according to CEFR and went through meticulous testing and approval process before its implementation. Therefore, the examinees can rest assured of the quality and applicability of the test. However, to ensure that the test is continuously reliable, suitable, correct, and complete according to the current educational progress, the test itself is required to be regularly assessed according to the latest CEFR measures and other recognised tests as quality assurance. Such assessment is important to retain and attract students to study at the university, especially working adults with high employment demands at stake. It is also crucial in maintaining the reputation of the university as an advanced and modern institution promoting open and distance learning as an effective education system.

Thus, the author conducted this research to assess STOU-EPT (E-Testing) and compare it to the proficiency tests which were in line with CEFR; of a local bilingual university, Srinakharinwirot University Standardized English Test (SWU-SET), and an international English proficiency standard; International English Language Testing System (IELTS). The involvement of SWU-SET and IELTS is crucial to evaluate the pertinence of STOU-EPT (E-Testing) within Thailand and internationally. The results from this research not only confirm the quality of the current testing system, but it can also serve as a reference to generate a new English proficiency test for other organisations and institutions that is reliable and globally accepted. Additionally, this research hoped to encourage the digitalisation of proficiency assessment to ease the process and eventually promote the acquirement of language proficiency certification.

2. Literature Review

The Common European Framework of Reference for Languages (CEFR) is the standard for evaluating linguistic proficiencies created by the European Union (EU). Its objective is to be the guidelines for learning, teaching, and evaluating second or foreign languages. In 2002, the EU required CEFR to be used for evaluating linguistic proficiencies within its provisions. Since then, CEFR gradually became a widely accepted standard in measuring individuals' linguistic proficiencies. The evaluation under CEFR can be divided into six levels: A1, A2, B1, B2, C1 and C2; whereby A1 represents basic English Language usage and C2 indicates complete proficiency. For Thailand, the Ministry of Education directly applied the CEFR standards in primary schools and secondary schools. The ministry also coordinated with the British Council to develop programmes for instructors and educational workers to train them on CEFR until they are competent enough to evaluate students' English proficiencies. At the higher education level, the policy does not restrict the institutions to implement CERF, however, they are required to use standards equivalent to CERF or other globally recognised standards to evaluate the proficiencies of their students (Higher Education Commission, 2016) in the aspects of listening, structure and reading.

Accordingly, STOU developed STOU-EPT (E-Testing), their personalised testing system to suit the needs of the university while still adhering to the ministry's policy and maintaining the universality of the certification when used outside of the university by referencing it to CEFR. The E-Testing consists of three parts: Part 1: Listening; Part 2: Structure; and Part 3: Reading. The examinees can log in by inputting their test schedule IDs, personal IDs and examinee card IDs shown in the e-tickets sent by e-mail used for applying for the test. After taking the test at the selected times on the selected dates, they can know the test results instantly. The test results have three levels: A2, B1 and B2. A2 is the English Basic User Level of Elementary English while B1 and B2 are the English Independent User Levels of Intermediate English (B1 Level) and Upper-Intermediate English (B2 Level). Concurrently, the Srinakharinwirot University (SWU) is another university in Thailand that implements an English proficiency standard with consistent criteria to CEFR, called SWU-SET; provided by the Language Centre of the International Colleges for Sustainability Studies of SU. It has 100 items evaluating listening skills, vocabulary, usage, structures, grammar, and reading skills (Language Centre of Srinakharinwirot University, 2019).

Other than CEFR, IELTS is another English proficiency test provided by a foreign organisation that is widely accepted in Thailand; co-administered by the British Council, IDP IELTS, and the Cambridge University Press and Assessment. The test has two forms: the paper-based test and the computer-based test. Both forms have similar items and contents; consisting of four parts: listening, reading, speaking and writing. The listening part allocates 30 minutes, the reading part allocates 60 minutes for four articles, and the writing part was allocated 60 minutes for two items. For speaking, the examinees have to be tested by native speakers in the forms of communications and opinions; inquiring about personal information and specified topics (IELTS, 2019). Both SWU-SET and IELTS are considered to be standardised tests consistent with CEFR, representing both the local assessment and international assessment which can evaluate the concurrent validity of STOU-EPT (E-Testing). This method was consistent with existing research on assessing the concurrent validity of linguistic tests, whereby the scores of different tests taken by the same examinees are compared to each other. If the scores from the studied tests had significant relationships with the standardised tests, it indicates that the tests had concurrent validity.

For example, Hoffman et al. (2011) researched the consistency of the concurrent validity and the structures of a speaking proficiency test for linguistically impaired school children. A total of 216 children with the impairment were evaluated with the Test of Language Development Primary, Third Edition (TOLD-P:3) (Newcomer & Hammill, 1997, as cited in LaVae M. Hoffman et al., 2011) and the Comprehensive Assessment of Spoken Language (CASL) (Carrow-Woolfolk, 1999, as cited in Hoffman et al., 2011) consecutively for their understanding of spoken language. The relationships and comparisons between both test scores were analysed in terms of the concurrent validity with Pearson's product-moment correlation coefficient. The consistency of the structures was checked through the confirmatory factor analysis. It was found that the scores from TOLD-P:3 and CASL had a significant relationship at the significance level of .001. and Pearson's product-moment correlation coefficient at 0.596. It was remarked that TOLD-P:3 had two components, while CASL had only one factor.

Cadorin et al. (2016) also conducted a similar study to confirm the concurrent validity of the self-rating scale named Self-Directed Learning – Italian version (SRSSDL-Ita) by testing it against a well-recognised test from Taipei; the Self-Directed Learning Instruments (SDLI). SRSSDL-Ita, developed in the Italian context, contains 40 items distributed into eight factors: Awareness, Attitudes, Motivation, Learning Strategies, Learning Methods, Interpersonal Skills, and Constructing Knowledge. Each item was rated using a five-point Likert scale ranging from 1 for never to 5 for always. The total score ranged from 60 to 300, whereby low scores indicate poor self-directed learning abilities and consequently denote that the student needs guidance in studying and vice versa. Instead, the SDLI consists of 20 question items categorised into four domains: Learning Motivation, Planning and Implementing, Self-monitoring, and Interpersonal Communication. The test also uses a five-point Likert scale, with a full score of 100. The research was participated by 428 nursing undergraduates in Italy and was analysed for concurrent validity through Pearson correlation. The results showed a 0.815 correlation, which confirmed that SRSSDL-Ita was concurrently valid to be used by Italian nursing students to achieve their learning goals.

Another recent research by Eisenberg et al. (2019) also performed the same method, whereby they investigated the concurrent validity of a speaking test and a disciplinary test for preschool children by comparing it against the Fluharty Preschool Speech and Language Screening Test (Fluharty-2) (Fluharty, 2001, as cited in Eisenberg et al., 2019). It was participated by 62 children and through Pearson's correlation, it was found that both tests had a significant relationship with each other at the significance level of 0.001 and Pearson's product-moment correlation coefficient of 0.722, signifying a concurrent validity. By reviewing the method applied in these studies, it was proven effective in assessing the concurrent validity of linguistic tests. Therefore, the researchers of this paper decided to analyse the concurrent validity of STOU-EPT (E-Testing) by examining the relationships between the linguistic proficiencies of the examinees according to the test and their scores in other well-established tests parallel to CEFR; in which the researchers chose IELTS and SWU-SET for highly reliable results. The results shall ascertain the current qualities of STOU-EPT (E-Testing) and if necessary, any improvement in terms of correctness, appropriateness, and standardisation can be further discussed and arranged.

3. Research Methodology

The present study imposed a quantitative design and adapted the existing method of assessing concurrent validity by having the same set of respondents participate in the studied test and a standardised test referred to by the studied test. However, for this paper, instead of comparing the studied test to only one test, three tests were included to increase the results' reliability. Further details are described in the following.

3.1. Participants

To fit the target population for this research, participation from STOU students who had completed STOU-EPT in 2019, selected using purposive sampling method. For more reliable data, the samples were also stratified according to their proficiency as assessed from STOU-EPT (E-Testing); A2 Level, B2 Level, and B1 Level. From the data on STOU-EPT takers, the researchers asked for volunteers with proficiency levels from A2 to B2 until the quota of 10 respondents at each level was filled. The research received a total participation of 30 respondents. The respondents were requested to take both SWU-SET and IELTS consecutively. However, one participant did not take SWU-SET and another two participants did not take IELTS. The data from respondents who did not fulfil any of the tests were removed from the analysis stage for each respective test, leaving 28 remaining respondents for IELTS and 29 for SWU-SET. The participation of the respondents in IELTS and SWU-SET was voluntary and sponsored by the research fund.

3.2. Data Collection

The data collection process started with the researchers choosing potential candidates from the list of the target population; the students who had taken the STOU-EPT (E-Testing) in 2019, before communicating with the candidates through phone calls to inquire about their interest in participating in the study. The process was executed until the quota of 10 participants for A2 Level, B2 Level, and B1 Level each was filled with consenting participants who were willing and convenient to take all CEFR, SWU-SET and IELTS for the study.

A database of the participants was generated after their details were collected through an electronic form sent online for details necessary to the research. With the details provided, the participants were then registered to sit for SWU-SET and IELTS on different dates; whereby the registration fees of all participants for both tests were paid to the institutes from the research fund. The results of the participants in both tests were collected through the second electronic form with attached test result files required for data processing. The test results built the research's secondary database.

The participants were informed of the objectives and overview of the research before asking for their consent. They were allowed to withdraw from the research at any time and for any reason, and their participation was ensured to be voluntary. Their personal information was kept confidential and was not included in any presentation of the results for anonymity and to preserve the reputation of the participants or any affiliated organisations.

3.3. Instrument

The research utilised two electronic surveys which were administered online for the data collection stage; the first survey was to collect details required for the registration in SWU-SET and IELTS and the second survey was to collect the results of the two tests from each participant. For data processing and analysis, IBM SPSS Statistics software was used to ensure accurate and valid results. The database was recorded in Microsoft Office Excel electronically.

3.4. Data Analysis

The secondary data were first processed with SPSS to obtain the frequency and percentage in each proficiency aspect to record the consistency of the participants' results in STOU-EPT (E-Testing), SWU-SET, and IELTS. Then, the concurrent validity of STOU-EPT (E-Testing) according to the levels of CEFR was generated through the relationship analysis of all three tests by using Kendall's Tau-b formula, a formula that analyse the relationships between the variables with an ordinal scale. The researchers grouped the data into sets of the same units; the proficiencies of the same examinees in the tests, consistent with the analysis method by Gibbons and Chakraborti (2014) for reliable results.

4. Findings and Discussion

The results of the analysis on the consistency of participants' proficiency levels and the concurrent validity of STOU-EPT (E-Testing), SWU-SET, and IELTS were summarised in this section.

4.1. The Consistency of the Skill Levels of the Examinees in STOU-EPT (E-Testing), SWU-SET and IELTS

Table 1 displays the comparison between the listening skill results for the participants in STOU-EPT (E-Testing) to SWU-SET and IELTS. It was found that all participants who obtained the A2 Level also obtained the A2 Level in the other two tests. However, 2 participants who got B2 in STOU-EPT (E-Testing) were evaluated as B1 in SWU-SET. The B1 Level is observed to be the least consistent. It was seen that 1 participant was evaluated as B2 in each IELTS and SWU-SET respectively, and a surprising number of participants got A2; 4 in IELTS and 7 in SWU-SET.

Table 1. The Consistency of Listening Skill Levels in STOU-EPT (E-Testing), IELTS, and SWU-SET

Level		A2		B1		B2		Total	
		IELTS	SWU-SET	IELTS	SWU-SET	IELTS	SWU-SET	IELTS	SWU-SET
STOU-EPT	A2	10 100.00%	10 100.00%	-	-	-	-	10 100.00%	10 100.00%
	B1	4 40.00%	7 70.00%	5 50.00%	2 20.00%	1 10.00%	1 10.00%	10 100.00%	10 100.00%
	B2	-	-	-	2 22.22%	8 100.00%	7 77.78%	8 100.00%	9 100.00%

Compared to the proficiency levels in STOU-EPT (E-Testing), IELTS resulted in a 140% increase for A2 (14 participants), 50% for B1 (5 participants), and 112.5% for B2 (9 participants). Meanwhile, SWU-SET presented 170% at the A2 level (17 participants), 40% at the B1 level (4 participants), and 77.78% at the B2 level (7 participants). Overall, the total percentage of inconsistent test results for the Listening Skill as displayed in Table 1 was 17.24% (5 results) in IELTS and 34.48% (10 results) in SWU-SET out of the total 29 results for each test, whereby only 3% in each test downgraded from the STOU-EPT (E-Testing) results.

The participants' proficiency levels in structuring from STOU-EPT (E-Testing), SWU-SET, and IELTS as shown in Table 2 were evaluated from questions in writing. It was found that all participants receiving the A2 Level in STOU-EPT (E-Testing) were still at A2 Level in SWU-SET and IELTS. Meanwhile, the participants who were at the B2 Level mostly received the higher result A2; 60% in IELTS and 80% in SWU-SET. Similar results were found for those at B2 Level, whereby three results were upgraded to B1: 1 in IELTS and 2 in SWU-SET.

Table 2. The Consistency of Structuring Skill Levels in STOU-EPT (E-Testing), IELTS and SWU-SET

Level		A2		B1		B2		Total	
		IELTS*	SWU-SET	IELTS*	SWU-SET	IELTS*	SWU-SET	IELTS*	SWU-SET
STOU-EPT	A2	10 100.00%	10 100.00%	-	-	-	-	10 100.00%	10 100.00%
	B1	6 60.00%	8 80.00%	4 40.00%	2 20.00%	-	-	10 100.00%	10 100.00%
	B2	-	-	1 10.00%	2 22.22%	7 77.78%	7 77.78%	8 100.00%	9 100.00%

*Participants' writing skill levels in IELTS

The IELTS portrayed 160% of A2 achievers (16 participants), 50% of B1 achievers (5 participants), and 87.5% of B2 (7 participants). The SWU-SET, however, resulted in 180% of A2 (18 participants), 40% of B1 (4 participants), and 77.78% of B2 (7 participants). The total inconsistency was 24.14% for IELTS and 34.48% for SWU-SET, whereby all inconsistencies were upgrades from the results in STOU-EPT (E-Testing).

Likewise, all participants who achieved A2 in STOU-EPT (E-Testing) for reading skill also received A2 results in the two tests as displayed in Table 3. Interestingly, the consistency varied significantly between IELTS and SWU-SET at the B1 Level. There were 2 participants who were upgraded to A2 in IELTS and 8 participants in SWU-SET. The results mostly showed consistency at the B2 Level, except for 2 participants who achieved the B1 Level in SWU-SET.

Table 3. The Consistency of Reading Skill Levels in STOU-EPT (E-Testing) IELTS, and SWU-SET

Level		A2		B1		B2		Total	
		IELTS	SWU-SET	IELTS	SWU-SET	IELTS	SWU-SET	IELTS	SWU-SET
STOU-EPT	A2	10 100.00%	10 100.00%	-	-	-	-	10 100.00%	10 100.00%
	B1	2 20.00%	8 80.00%	8 80.00%	2 20.00%	-	-	10 100.00%	10 100.00%
	B2	-	-	-	2 22.22%	8 100.00%	7 77.78%	8 100.00%	9 100.00%

Under IELTS, a total of 12 participants achieved the A2 Level (100%), 8 participants achieved the B1 Level (80%), and 8 participants achieved the B2 Level (100%). Meanwhile, a significant number of 18 participants got A2 under SWU-SET (100%), while only 4 got B1 (40%) and 7 got B2 (77.78%). Inclusively, the proficiency levels of reading skills portrayed the least inconsistency between STOU-EPT (E-Testing) and IELTS with only 6.90%. Whereas the inconsistency between STOU-EPT (E-Testing) maintained the same in all three skills at 34.48%.

4.2. The Evaluations of the Concurrent Validity of STOU-EPT (E-Testing), IELTS and SWU-SET in the Aspects of Listening, Structuring, and Reading

The researchers analysed the data to determine the concurrent validity of STOU-EPT by using Kendall's Tau-b formula. STOU-EPT (E-Testing) was tested against IELTS and SWU-SET each according to the English language proficiency skills in the aspects of Listening, Structuring, and Reading. The three skills in each test were set as the variables in the calculation. The results of the calculation were summarised in Table 4. The numerical values in the table indicate the strength of the relationship between the variables, whereby any value the same as or bigger than 0.700 implied a strong relationship.

Table 4. The Analysis of the Relationships of the CEFR Levels of STOU-EPT (E-Testing), IELTS and SWU-SET

		Listening		Structuring		Reading	
		IELTS	SWU-SET	IELTS (Writing)	SWU-SET	IELTS	SWU-SET
Listening	STOU-EPT	0.757**	0.700**				
Structuring				0.804**	0.779**		
Reading						0.890**	0.812**

** statistically related at the significance level of 0.01.

As displayed in the table, all variables in STOU-EPT (E-Testing) scored at the value of or bigger than 0.700 when tested against IELTS and SWU-SET with Kendall's Tau-b formula. The value indicated that the proficiency in the Listening, Structuring, and Reading aspects in STOU-EPT (E-Testing) was significantly related to the same aspects in the other two sets. The scores ranged from 0.700 at the lowest and 0.890 at the highest. The highest scoring aspect was Reading under IELTS and the lowest was Listening under SWU-SET. The values averaged 0.790, indicating medium overall strength.

4.3. Discussion

Based on the results as displayed in Table 1, Table 2, and Table 3, the A2 level had the highest inconsistency with the percentages of 140%, 160%, and 120% for A2, B1, and B2 respectively in IELTS; and 170%, 180%, and 180% for A2, B1, and B2 respectively in SWU-SET. On average, 40% of B1 STOU students achieved A2 in IELTS and 76.67% of B1 STOU-EPT (E-Testing) takers got A2 when taking SWU-SET. Conversely, those who achieved A2 in STOU-EPT (E-Testing) also achieved the same results in the other two tests at 100%. These percentages implied that STOU-EPT (E-Testing) has a higher standard for the A2 level than the two tests. While lower proficiency achievers may be disadvantaged by this fact, those who got A2 can rest assured that they belonged there.

Meanwhile, due to the high percentage of B1 achievers who were pulled into the A2 level in IELTS and SWU-SET, the inconsistency at the B1 level ranged the farthest of all levels. Furthermore, 10% of B1 achievers lowered their level to B2 in both IELTS and SWU-SET for listening skills. The findings may indicate that a revision should be made to STOU-EPT (E-Testing) for the Listening skill test to ensure whether the disparity was sourced from individual hindrances or the assessment itself. If necessary, more training on Listening skills should be emphasised and improvement in the assessment should be made. However, for the other two skills, the inconsistencies were only due to the participants achieving the upper level of A2, suggesting a similar suggestion as the one on A2.

Among the three levels, the B2 level showed the most consistency, whereby the percentage does not exceed the range of 77.78% and 112.5%. The percentage was at 100% for IELTS in both Listening and Reading Skills. All differences were due to up-levelling since B2 was already at the lowest level. While the consistency may be the result of the incompetence of the participants in the English language which will reveal the same level in any tests with the same purpose, the up-levelling signified that some participants may have improved themselves in between the duration they received the STOU-EPT (E-Testing) result and the date they sat for IELTS and SWU-SET. It may also reiterate the suggestion above that STOU-EPT (E-Testing) was slightly more challenging than the other two tests.

Otherwise, there might also be errors in the thresholds of the scores in STOU-EPT (E-Testing) because the threshold range was relatively narrow and the score range of the B1 Level was quite wide. Therefore, the examinees at B1 Level obtained the scores at A2 Level in the other standardised tests. This was consistent with EF (2019) who stated that standard setting was a process linking the opinions of the experts and the score interpretation for setting the thresholds for the levels of scores. A discussion with the experts crafting the assessments may provide more insight into the process. The possibility of human errors in determining the thresholds was also possible since the process was dependent on their discretion. However, any amendments to the existing assessment should be made by experts for each aspect for a higher success rate.

Finally, based on the analysis using Kendall's Tau-b formula, the proficiency levels in STOU-EPT (E-Testing) were significantly and statistically related to IELTS and SWU-SET according to the CEFR levels in all aspects. The value of correlation strength was at 0.700 and higher in all aspects. While 0.700 indicates a good strength, the value was at the bottom of acceptability. None of the aspects reached the desirable 0.900, which implied a very good correlation strength. Regardless, such value was also expected since each test was personalised to a specific purpose; in which STOU-EPT (E-Testing) was tailored to the learning structure for the language in STOU. Thus, it can be ultimately determined that STOU-EPT (E-Testing) has a valid standard of examination according to CEFR criteria; in line with IELTS and SWU-SET. Further study and research on determining the cut-off score for each proficiency level to provide more consistent scoring is recommended.

5. Conclusion

According to the analysis of the results, it was deduced that STOU-EPT (E-Testing) was concurrently valid when compared to IELTS and SWU-SET; two English language proficiency assessments in line with CEFR. Measured through Kendall's Tau-b formula, the proficiency levels in STOU-EPT (E-Testing) were found to have significant relationships with IELTS and SWU-SET in all aspects. Thus, the findings are ensured valid since the research was consistent with those by Hoffman et al. (2011), and Eisenberg et al. (2019) who evaluated the concurrent validity of the linguistic tests developed by them against the standardised tests through Pearson's correlation coefficient. Moreover, when analysed with a different formula, they found that the relationships between the test scores from both tests maintained significant, indicating the reliability of the results. Data using a different formula is measured for consistency and reliability of this research. This finding confirmed that the results from STOU-EPT (E-Testing) were reliable, accurate, and practical. Examinees who had sat for STOU-EPT (E-Testing) should be confident and proud of their English language proficiency level as reported by the results. Not only the online version of the test is as reliable as the paper-based test, but the results can also be viewed immediately after the completion of the test. As such, the examinees can immediately start working on the improvement of their aptitudes while also using STOU-EPT (E-Testing) as a benchmark for their learning goals. Moreover, since it was consistent with other assessments meeting the criteria of CEFR, it should also be able to be used outside of the university such as for furthering the study, work application, and international travelling.

Funding: This research study was supported by the distance education fund from the Research and Development Center of Sukhothai Thammathirat Open University in 2018.

References

- Cadorin, L., Su-Fen Cheng, & Palese, A. (2016). Concurrent validity of self-rating scale of self-directed learning and self-directed learning instrument among Italian nursing students. *BMC Nursing, 15*, 1-10.
- Council of Europe. (2021). *Common European framework of reference for languages: Learning, teaching, assessment*, Cambridge University Press.
- EF. (2019). *EFSET technical background report*. EFSET. Retrieved April 24, 2019, from <https://www.efset.org/th/english-score/cefr/>
- Eisenberg, S., Victorino, K., & Murray, S. (2019). Concurrent validity of the fluharty preschool speech and language screening test—second edition at age 3: Comparison with four diagnostic measures. *Language, Speech & Hearing Services in Schools, 50*(4), 673–682. https://doi.org/10.1044/2019_LSHSS-18-0099
- Gibbons, J. D., & Chakraborti, S. (2014). *Nonparametric statistical inference*. CRC press.
- Hoffman, L. M., Loeb, D. F., Brandel, J., & Gillam, R. B. (2011). Concurrent and construct validity of oral language measures with school-age children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 54*(6), 1597. [https://doi.org/10.1044/1092-4388\(2011/10-0213\)](https://doi.org/10.1044/1092-4388(2011/10-0213))
- Higher Education Commission. (2016). *The Announcement of the Higher Education Commission on the English Standard Improvements in the Institutes at the Higher Education Level*.
- IELTS. (2019). *IELTS introduction*. Retrieved August 11, 2019, from <https://www.ielts.org/what-is-ielts/ielts-introduction>
- Jiraro, S., & Angsuchoti, S. (2021). Validating Sukhothai Thammathirat Open University's English Proficiency Test (E-Testing) of with the computer system under the Common European Framework of Reference for Languages (CEFR). *Journal of Social Sciences in Measurement Evaluation Statistics and Research, 2*(1), 27-35.

- Language Center Srinakharinwirot University. (2019). *SWU-SET Application*. Retrieved August 24, 2019, from <https://sites.google.com/site/languagecentreswu/home/swu-set-application>
- Mohamad, U., Narul, F. & Abd Aziz, M. S. (2022). The appropriacy and applicability of english assessment against CEFR global scale: Teachers' judgement. 3L: *Southeast Asian Journal of English Language Studies*, 26(3), 53-65.
- Peng, C., Liu, J. & Cai, H. (2022). Aligning China's standards of english language ability with the common European framework of reference for languages. *Asia-Pacific Edu Res*, 31, 667–677.