

## การเทียบผลคะแนนแบบทดสอบ K-STEP กับแบบทดสอบ TOEIC โดยอ้างอิงจากกรอบมาตรฐานความสามารถทางภาษา ของประเทศในกลุ่มสหภาพยุโรป (CEFR)

งามทิพย์ วิมลเกษม<sup>1</sup>ทิวาพร คงสม<sup>2</sup>สุภาลักษณ์ นครศรี<sup>2</sup>พงษ์พัฒน์ พันธนะศิริ<sup>3</sup>ฉันทพร แสงวิรัช<sup>3</sup>ณัฐลิกา มหาพูนทอง<sup>3</sup>กรกฎ กำแก้ว<sup>3</sup>วัลย์กร เจริญสุข<sup>3</sup>พรสวรรค์ ศุภศรี<sup>3</sup><sup>2</sup> ผู้นิพนธ์ประสานงาน โทรศัพท์ 081-8428661 อีเมล : tkongsom@hotmail.com

รับเมื่อ 21 มกราคม 2563 ตอรับเมื่อ 8 มิถุนายน 2563 DOI:10.14416/j.faa.2020.25.008

### บทคัดย่อ

แบบทดสอบสมิทิภาพทางภาษาอังกฤษของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ หรือ แบบทดสอบ K-STEP เป็นแบบทดสอบสมิทิภาพทางภาษาอังกฤษจัดทำโดยภาควิชาภาษา คณะศิลปศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือเพื่อใช้วัดสมิทิภาพทางภาษาอังกฤษของนักศึกษาทั้งระดับปริญญาตรีและระดับบัณฑิตศึกษา งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบทดสอบมาตรฐานสมิทิภาพทางภาษาอังกฤษโดยอ้างอิงจากกรอบมาตรฐานความสามารถภาษาของประเทศในกลุ่มสหภาพยุโรป (CEFR) และเพื่อเทียบหาความสัมพันธ์ระหว่างผลคะแนนแบบทดสอบ K-STEP กับแบบทดสอบ TOEIC โดยมีขั้นตอนการดำเนินการวิจัยเริ่มจากการสร้างแบบทดสอบ K-STEP และประเมินข้อสอบ 100 ข้อ ตามลักษณะเฉพาะของข้อสอบ โดยคณาจารย์ผู้มีประสบการณ์ในการสอนภาษาอังกฤษ ข้อสอบที่ได้รับการประเมินเพื่อหาความเที่ยงตรงโดยผู้เชี่ยวชาญจำนวน 5 คน ได้นำไปทดลองเพื่อหาค่าความเชื่อมั่นกับนักศึกษาจำนวน 178 คน ในส่วนการหาค่าความสัมพันธ์ระหว่างคะแนนแบบทดสอบนั้นนักศึกษอาสาสมัครจำนวน 105 คนได้ทำแบบทดสอบ K-STEP กับแบบทดสอบ TOEIC ผลการวิจัยสรุปได้ว่าผลคะแนนของแบบทดสอบ K-STEP กับแบบทดสอบ TOEIC มีความสัมพันธ์เชิงบวกอย่างมาก (Pearson Correlation = .877) ผลของการวิเคราะห์แบบการถดถอยเชิงเส้นอย่างง่ายแสดงว่าผลคะแนนแบบทดสอบ K-STEP สามารถใช้ทำนายผลของคะแนน TOEIC ได้ (R-coefficient = .877) นอกจากนั้นผลการวิจัยสามารถใช้ในการเปรียบเทียบคะแนนที่ได้จากแบบทดสอบ K-STEP กับแบบทดสอบ TOEIC เพื่อใช้อธิบายความสามารถทางภาษาอังกฤษของผู้เข้าสอบข้อสอบทั้ง 2 ชนิดดังกล่าวได้

**คำสำคัญ :** แบบทดสอบสมิทิภาพทางภาษาอังกฤษ การเทียบผลคะแนน แบบทดสอบ K-STEP แบบทดสอบ TOEIC

<sup>1</sup> ผู้ช่วยศาสตราจารย์ ดร. ชำระการบ้านานู มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

<sup>2</sup> ผู้ช่วยศาสตราจารย์ ดร. ประจักษ์ภาควิชาภาษา คณะศิลปศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

<sup>3</sup> อาจารย์ประจำภาควิชาภาษา คณะศิลปศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

## Mapping the K-StEP Test Scores to the TOEIC Test Scores Based on the Common European Framework of Reference (CEFR)

Ngamthip Wimolkasem<sup>1</sup>  
Tiwaporn Kongsom<sup>2</sup>  
Supalak Nakhornsri<sup>2</sup>  
Pongpat Pantanahiran<sup>3</sup>  
Thanyaporn Sangwirach<sup>3</sup>  
Challika Mahaphuntong<sup>3</sup>  
Korakot Kamkaew<sup>3</sup>  
Valaikorn Charoensuk<sup>3</sup>  
Ponsawan Suphasri<sup>3</sup>

<sup>2</sup> Corresponding Author, Tel. 081-8428661, E-mail : tkongsom@hotmail.com  
Received 21 January 2020; Accepted 8 June 2020

### Abstract

The King Mongkut's University of Technology North Bangkok Standardized English Proficiency Test or K-StEP Test is an English proficiency test created by the Department of Languages, Faculty of Applied Arts, King Mongkut's University of Technology North Bangkok (KMUTNB). The test assesses the English proficiency of KMUTNB undergraduate and graduate students. The study aims to develop an organized standard English proficiency test based on the Common European Framework of Reference (CEFR) for Languages and to find a correlation between the K-StEP Test scores and the TOEIC Test scores. In the study, several procedures were implemented to strengthen validity and reliability. A team of experienced English lecturers joined the test development process and helped to validate 100 items of the K-StEP Test based on the test specifications. The test was validated by 5 experts while reliability was gained from the pilot study with 178 students. In order to explore the correlation, 105 volunteer students participated in undertaking the K-StEP Test and the TOEIC Test. The findings revealed a highly positive correlation between the K-StEP Test scores and the TOEIC Test scores at .877 by Pearson Correlation. The results of the Simple Regression Analysis revealed high equating validity between the two tests ( $R\text{-coefficient}=.877$ ). Moreover, the results can be used to compare the test scores of the two tests and define the English ability of the test takers.

**Keywords :** English proficiency test, Mapping scores, K-StEP Test, TOEIC Test

<sup>1</sup> Emeritus Assistant Professor Dr., Department of Languages, Faculty of Applied Arts, King Mongkut's University of Technology North Bangkok

<sup>2</sup> Assistant Professor Dr., Department of Languages, Faculty of Applied Arts, King Mongkut's University of Technology North Bangkok

<sup>3</sup> Lecturer, Department of Languages, Faculty of Applied Arts, King Mongkut's University of Technology North Bangkok

## 1. Introduction

There are several reasons why there has been an urgent need to enhance the English proficiency of Thai students. Firstly, the English language is the global tool for communication and for academic study. It is the primary language used when searching for knowledge, for communication amongst countries, and global business, and international relations. Secondly, due to the rapid change in new technology, English is needed as a medium to transfer the knowledge to the implementation process. Thirdly, Thai students have a relatively low level of English language proficiency.

According to the English Proficiency Index based on the English test scores of 1.7 million test takers in 88 countries, Thailand is ranked 64<sup>th</sup> with an average score of 48.54% defined as low proficiency (Education First, 2018). Other countries in ASEAN, fared much better, for example, Singapore is ranked 3<sup>rd</sup> (68.63%), Philippines is ranked 14<sup>th</sup> (61.84%), Vietnam is ranked 41<sup>st</sup> (53.12%), and Indonesia is ranked 51<sup>st</sup> (51.88%). Moreover, the results of the Ordinary National Education Test (O-NET) in English in 2017 were very disappointing with an average English language score of 28.31% as shown in Table 1 below.

**Table 1** O-NET scores of Mattayom 6 students in 2016 and 2017 (National Institute of Education Testing Services, 2018)

Subjects	Year 2016	Year 2017
Thais	52.29	49.25
Social studies	35.89	34.70
Sciences	31.61	29.37
English	27.76	28.31
Mathematics	24.88	24.53

In 2016, the Office of the Higher Education Commission (OHEC) decided on a policy to raise the levels of English proficiency of Thai students in higher education. The main objective was that English should equip students to search for new knowledge, to keep up with the rapid global change, to support the ASEAN Economic Community (AEC), and to encourage the country's competitiveness. The policy included 5 mission statements as follows:

1. The university must define the policy and set targets in elevating the English proficiency level at every educational level. This will be the guideline in developing the students and graduates' readiness in both academic and professional fields.

2. The university must have a definite plan to follow by setting an indicator for English proficiency assessment.

3. The university must take action to improve the learning and teaching process in English to reach the required targets.

4. The university must provide activities, materials, media instruction, and environment to motivate students' self-development and awareness in using English.

5. The university should assess every student's English proficiency level either through in-house assessment or by use of a standard English proficiency test by mapping with the levels of the

Common European Framework of Reference (CEFR) for Languages. This can provide information of individual student English proficiency. The English proficiency certificate can be attached to or be part of the academic transcript starting from the academic year 2019.

To enhance the English abilities of Thai students to enable them to perform efficiently and effectively in English, the implementation of the Common European Framework of Reference (CEFR) for Languages has been used as a standard tool and set the target at all levels of education by the Ministry of Education as presented in Table 2.

**Table 2** CEFR target level by the Ministry of Education (Office of the Basic Education Commission, 2018)

Level of education	CEFR level
Primary level (Prathom 1-6)	A1
Secondary level (Mattayom 1-3)	A2
High School level (Mattayom 4-6)	B1
Bachelor Degree level	B2

To summarize, educational policy has brought about changes in the teaching, learning, and testing of English in Thailand. Although the National Education Act is well-meaning in its goals aiming at improving Thais' English skills in reality, it fails to achieve the goal as can be seen in the disappointing results of the National Exam and Thailand's poor ranking in English proficiency.

Language testing is an integrated and necessary part of any language program. It is a crucial tool for language proficiency assessment for learners' advancement and program evaluation (Norris, 2006). Regarding raising standards and goals in English proficiency, several Thai universities have developed their own standards of English testing, examples being the Chulalongkorn University Test of English Proficiency (CU-TEP), and the Thammasat University General English Test (TU-GET). King Mongkut's University of Technology North Bangkok has also developed an English proficiency test based on performance and CEFR namely King Mongkut's University of Technology North Bangkok Standardized English Proficiency Test (K-StEP Test) and has used the test for decades. The K-StEP Test was originally

adopted to evaluate the English proficiency of King Mongkut's University of Technology North Bangkok (hereafter called KMUTNB) graduate students. It has been gradually accepted university-wide at every level from bachelor degree to graduate level and as the criteria for university personal recruitment. In response to the high demand for a standard English proficiency exam for various purposes including admission, certification, and exit examination, it was necessary to elevate the K-StEP Test as a national standard English test by finding its correlation with a renowned English global test. The Test of English for International Communication (TOEIC Test) has been selected for the correlation of this study because it can be used worldwide as a reference for job recruitment and promotion as well as academic purposes. In addition, the current study aimed to find the predictive validity of the K-StEP Test scores and the TOEIC Test scores because the predictive validity can reveal the extent to which the test can predict some future or desired outcome. Therefore, if test takers know their K-StEP Test scores, it is possible that they will know their TOEIC Test scores. In order to find the specific

meanings to describe what the test takers can do after they know their K-STEP Test scores and TOEIC Test scores, the ability bands or descriptors of CEFR levels were used to identify the English ability of the test takers of the K-STEP Test and the TOEIC Test at each level. The descriptors of CEFR levels were commonly used because they are internationally recognized as the common framework of language use to raise standards in many standard English proficiency tests.

Based on the problems and the requirements above, this study has been designed to develop an organized standard English proficiency test based on the Common European Framework of Reference (CEFR) descriptors namely the K-STEP Test, and then to find the relationship between the K-STEP Test scores and the TOEIC Test scores.

### 1.1 Objectives of the study

1.1.1 To develop an organized standard English proficiency test based on the Common European Framework of Reference (CEFR) for Languages and to find the correlation between the K-STEP Test scores and the TOEIC Test scores.

1.1.2 To examine whether the K-STEP Test scores can predict the TOEIC Test scores.

1.1.3 To find the ability bands for defining English proficiency obtained from the K-STEP Test scores.

### 1.2 Research questions

1.2.1 What are the concurrent validity values of the K-STEP Test scores and the TOEIC Test scores?

1.2.2 To what extent can the K-STEP Test scores predict the TOEIC Test scores?

1.2.3 What are the ability bands for describing English levels obtained from the K-STEP Test scores?

### 1.3 Research hypothesis

Hypothesis 1: There is a significant relationship between the K-STEP Test scores and the TOEIC Test scores at the .05 level. ( $H_1 : r_{xy} \neq 0$ ).

Hypothesis 2: The K-STEP Test scores can significantly predict the TOEIC Test scores at the .05 level. ( $H_2 : B \neq 0$ ).

## 2. Review of Related Literature

The review of literature and related studies is provided to examine the theories underlying this study.

### 2.1 The Council of Europe Framework of Reference and Language Education

The Council of Europe Framework of Reference and Language Education is an intergovernmental cooperation in the area of languages and education under the Language Policy Division of the Council of Europe. It was formally called the Modern Languages Section that can be traced back for almost 5 decades (Martyniuk, 2010). The Council of Europe was founded in 1949 with 10 members as one of the measures after the Second World War to increase international understanding and avoid future war. With 47 state members today, the Council of Europe is responsible for the European court of justice and continues to concern itself with the promotion of international understanding through the learning of languages, history, and culture of other countries (North, 2014).

The Council of Europe's Language Policy Division in Strasbourg is currently concerned with the languages of education particularly with the integration of migrants and their children. The aims of the Council of Europe's Language Policy are to achieve greater unity among its members and to pursue this aim by adopting common action in the fields of culture, education, and social sciences (Council of Europe, 2020).

The Common European Framework of Reference (CEFR) for Languages also provides a basis for the elaboration of language syllabi, curriculum guidelines, examinations, textbooks, etc. as it is divided into 6 levels and 3 bands: A1- A2 (Basic User), B1-B2 (Independent User), C1-C2 (Proficient

User). The descriptors of each level are described in a very comprehensive way what language learners have to learn and to do in order to use language for communication. Moreover, the knowledge and language skills should be developed to practice communication in each level. The description also covers the cultural context as well as linguistics competence. Since the descriptors of CEFR can define levels of proficiency which allows learners' progress to be measured at each stage of learning, a number of standard English proficiency tests i.e. TOEFL, IELTS, TOEIC and APTIS have mapped their tests' minimum scores corresponding to the six CEFR levels of language proficiency (A1-C2).

## 2.2 Studies on mapping the tests to the Common European Framework of Reference (CEFR)

Tannenbaum and Wylie (2015) conducted a study to provide a cut score of the TOEIC Test based on CEFR descriptors. The purpose of this study was to recommend minimum scores for each CEFR level that corresponds to the TOEIC Test in order to set a certain category. The panel consisted of 22 experts that were English language instructors, administrators, or directors of language programs and language testing from 10 different European countries. They were asked to recommend the cut score for each of the 6 CEFR levels. It was possible that not all experts would make the same decision as to the cut score. Therefore, a criterion of at least 67% or 2/3 of the 22 experts was the agreed level. The table below presents the results of the recommended minimum score for each test at each CEFR proficiency level.

**Table 3** Cut score for each skill of TOEIC Test compared to CEFR levels (Tannenbaum and Wylie, 2015)

Test/Test Section	Total Score Scale Range	Minimum Score				
		A1	A2	B1	B2	C1
TOEIC® Listening and Reading Total <sup>1</sup>						
TOEIC® Listening	5-495	60	110	275	400	490
TOEIC® Reading	5-495	60	115	275	385	455 <sup>4</sup>
TOEIC® Speaking and Writing <sup>2</sup>						
TOEIC® Speaking	0-200	50	90	120	160	180 <sup>5</sup>
TOEIC® Writing	0-200	30	70	120	150	180 <sup>5</sup>
TOEIC Bridge™ <sup>3</sup>						
TOEIC Bridge™ Listening	10-90	46	64	84		
TOEIC Bridge™ Reading	10-90	46	70	86		

<sup>1</sup> TOEIC® Listening and Reading scores are reported in 5-point increments.

<sup>2</sup> TOEIC® Speaking and Writing scores are reported in 10-point increments. No total score is reported.

<sup>3</sup> TOEIC Bridge™ scores are reported in 2-point increments. No total score is reported.

<sup>4</sup> TOEIC® Reading C1 minimum score is based on 45 percent of the panelists.

<sup>5</sup> TOEIC® Speaking and Writing scores were adjusted from the recommended study values.

In 2017, a group of English experts in Thailand developed the 10 level-framework of English for Thailand based on the CEFR (Hiranburana et al., 2017). The study involved 2 main parts, i.e. the drawing of the descriptors of the Framework of Reference for English Language Education in Thailand (FRELE-TH) and the public hearing. The discussion focused on the use of the framework for academic and professional qualification in Thailand. The study employed a qualitative and descriptive research, based on the analysis and adaptation of the CEFR. The results are very practical and useful for English teachers to employ or adapt for both curriculum development and test construction by offering two -scale types to describe English proficiency levels, i.e. global and illustrative scales. The global scales include an overview of the language proficiency at all levels while the illustrative scales comprise three aspects: communicative activities, communication strategies, and communicative language competence. In addition, the FRELE-TH includes the same structures and components of the CEFR, which offer language functions, discourse markers, topics, vocabulary range, grammar, micro skills and word family list. The FRELE-TH is divided into 10 levels and 3 bands: A1, A1+, A2, and A2+ (Basic User), B1, B1+, B2, and B2+(Independent User), C1-C2 (Proficient User). All the 10 levels present “can do” descriptors to define the actual use of English in communication. Apart from the academic purpose, the framework can also be used as a reference for job recruitment and job promotion.

### **2.3 Studies on the comparison between other English proficiency test scores and the TOEIC Test scores**

Hirai (2008) conducted a study to find the test score correlation between the STEP BULATS writing and the TOEIC Test scores. The participants were 559 Japanese businesspeople that took both the STEP BULATS Writing Test and the TOEIC Test

between September 2004 and December 2007. The results showed the STEP BULATS writing scores correlated moderately with the TOEIC Test scores, with a correlation coefficient of .69 for all score range. While the correlation coefficient for the upper-level group of participants with TOEIC scores of 800 and above was at .46, 50.4% of the 349 test takers that took the STEP BULATS Writing Test failed to perform the business English writing skills expected of competent international businesspeople. The study then suggested that the relatively low performance and scores in the STEP BULATS Writing Test were primarily because of the deficiency of exposure to business practice and vocabulary.

In Thailand, several studies conducted the comparison studies on their English proficiency tests and the TOEIC Test. Anuyahongsa (2015) conducted a study to compare the efficiency between the Thai-Nichi Institute of Technology (TNI) English Ability Test and the TOEIC Test. The participants were 20 TNI undergraduate students. The stratified random sampling technique was used to select the participants that enrolled in reading and writing skills in English 1 course. The data were collected by the TNI English ability test, 100 items for listening skill and 100 items for reading skill, and the TOEIC Test. Mean scores, standard deviation, t-test, and Pearson Correlation were used to analyze the collected data. The findings showed that the undergraduate students' competence in taking the TNI English ability test and the TOEIC Test was at moderate level. The findings also showed that the test scores of both TNI English ability test and TOEIC Test were closely related and the efficiency of the TNI English ability test was related to that of the TOEIC Test. In addition, the correlation between both tests was at very high level (.992).

Chanchusakun et al. (2017) conducted a comparison study on the quality of equating English scores between the SEPT and the TOEIC Test. The sample of this study included 40 students that

studied in bachelor of business administration in hotel management, Silapakorn University International College. The research instruments used in the study were the Silapakorn English Proficiency Test (SEPT) and the TOEIC Test. The findings of equating with regression method revealed that the score of the TOEIC Test can be equated with that of SEPT at the significant level .05.

In addition, Jaturapitakkul and Watson Todd (2018) wrote a report on mapping the Test of English for Thai Engineers and Technologists (TETET) with the Common European Framework of Reference (CEFR) for Languages and comparing the TETET scores with those of other English proficiency tests such as TOEIC, IELTS and TOEFL. The developed TETET is used to test English language proficiency of the test takers in four skills: reading, writing, listening and speaking. The content of the test is based on the situations that Thai engineers and technologists are required to use English. In order to find the test validation and compare the TETET scores with the TOEIC Test, the study was conducted with 124 test takers who took both the TETET and the TOEIC Test. The findings showed that there was a correlation between the TETET scores and the TOEIC Test scores (.883). This means that the TETET and the TOEIC Test were very strongly correlated. In order to find the relationship between the TETET band levels and the six CEFR levels, the study was conducted by mapping and identifying the minimum scores or levels to see if the TETET corresponded to each CEFR level. The panel of 10 experts was asked to assess each individual TETET test item based on the CEFR levels and descriptors. Finally, the cut-off scores were reported based on each CEFR level.

A number of studies mentioned above facilitate the description and guidelines of test construction in relation to the CEFR standard. This reflects the main purpose of this study that is to

find the relationship between the K-StEP Test scores and the TOEIC Test scores based on the CEFR levels. It is hoped that the results of this study will support the standard of the in-house test and help the university target student achievement in English proficiency.

### 3. Research Methodology

#### 3.1 Participants

The participants in this study can be divided into 2 main groups, i.e. the group of test writers and experts and the group of student participants.

##### 1. Test writers and experts

In order to develop the test, 7 EFL lecturers who had either masters' degrees or doctorates in English language teaching and learning took part in writing the K-StEP test items. Then another 5 experts with the same qualifications as the 7 test writers evaluated the constructed test items by using Item Objective Congruence (IOC) form and another 2 experts who were trained and had experience in selecting the K-StEP test items selected the test items based on the K-StEP test specification. All of the experts are EFL lecturers that had either masters' degrees or doctorates in English language teaching and learning. Their experience of teaching English is at least 5 years.

##### 2. Student participants

a. The student participants in the pilot study included 178 students (92 males and 86 females) from the Faculties of Engineering, Applied Sciences, Information Technology, Technical Education, and Industrial Technology.

b. Another group of the student participants was asked to participate in the main study. They consisted of 105 students from the above faculties that were willing to take the 2 tests, i.e. K-StEP Test and TOEIC Test in the same week.

#### 3.2 Research procedures

The study was developed through 4 main phases: Phase 1: The researchers studied the related



document and research, set up a workshop on CEFR for the test writers, and wrote the test specification; Phase 2: Seven test writers developed and wrote 150 K-StEP test items based on the K-StEP Test specification, which included English in communicative and academic contexts at various levels based on CEFR A2, B1, and B2. Then a native speaker of English proofread and edited all the developed test items. After that 5 experts evaluated the IOC form, and another 2 experts selected 100 test items based on the K-StEP test specification,

which consisted of 30 items for listening, 30 items for reading, and 40 items for grammar usage and structure; Phase 3: The constructed K-StEP test items were piloted with 178 student participants to find the reliability and quality of the test; and Phase 4: Another 105 students took the final 100 K-StEP test items and the TOEIC Test in the same week and then the students' scores of both tests were analyzed to map and find the cut-off score of the two tests. The research methodology and procedure are presented in Figure 1 below.

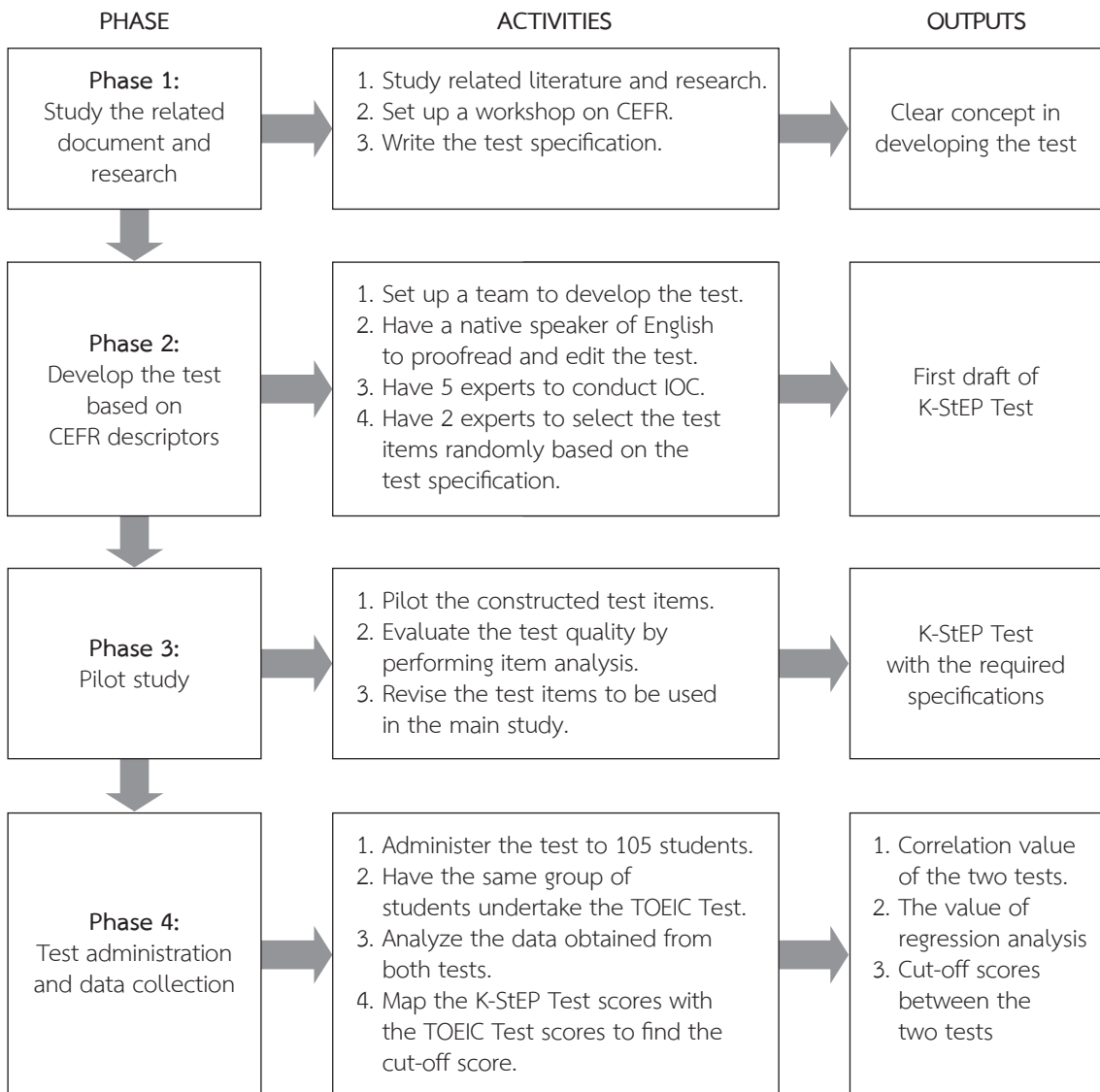


Figure 1 Research methodology and procedure

### 3.3 Research Instrument

The research instrument used in this study is the K-StEP Test. The K-StEP Test was developed in the following steps:

#### 1. The construction of the K-StEP Test

The test was developed according to the test specification as shown in Table 4 below.

**Table 4** The specification of K-StEP Test

Purposes of the test	Description
Types of test takers	Undergraduate and graduate students of KMUTNB. They are non-native speakers of the target language.
Test level	The difficulty of the test is set as A2, B1, and B2 based on CEFR descriptors.
Number of test items	The test is divided into 3 parts of 100 items of which 30 items were in a listening part, 30 items in a reading part, and 40 items concerning grammar usage and structure. All items are 4-option multiple choices. The context included 3 areas, i.e. a general setting for daily life, campus life, and academic context in various fields.
Marks	Computational checking with one score for each item
Time allocation	Three hours

The K-StEP Test has been designed for both graduate and undergraduate students that are non-native speakers of English. English proficiency is one of the requirements for the graduation of KMUTNB students from both levels. The test evaluates English proficiency in the receptive skills of listening, reading, and grammar usage and structure based on CEFR levels of A2, B1, and B2. The structure and specification of the K-StEP Test are described below.

The listening section measures the test takers' ability to understand spoken English in communicative and academic contexts at various levels based on CEFR levels of A2, B1, and B2. The contents include personal particulars and family information, shopping, employment (A2), main idea, school, work and leisure (B1), technical discussion, and specialization (B2). The section consists of 3 parts, i.e. short dialogues, long dialogues, monologues and lectures with 30 multiple choice questions.

The reading section measures the test takers' ability to understand the range of factual, descriptive, and inferencing of texts and paragraphs. The content and vocabulary range are in accordance with the CEFR levels of A2, B1, and B2. The questions are to find the main ideas, major details, important details including vocabulary in contexts and pronoun references. There are 3 reading passages with 30 questions, each is worth 1 point.

The grammar usage and structure part assesses the ability of test takers on sentence structures and grammar usage as an indirect writing ability assessment. The grammar points and vocabulary range tested in the grammar usage and structure part are based on the CEFR levels of A2, B1, and B2. The first part comprises sentence completion of 15 items. The second part included defining grammatical mistakes of 15 items. The test takers were asked to choose the underlined parts

that are ungrammatical. The last part is the cloze-test of 10 items to complete phrases using the most appropriate vocabulary and the correct grammatical forms.

2. To ensure the reliability and validity of the test, the following processes were introduced.

- 150 test items were written by the test construction team that had attended the CEFR workshop.

- All the test items were edited by a native speaker of English in order to check the correctness and suitable language use.

- All the test items were validated by 5 experts for the IOC. The item with the IOC index lower than 0.5 was either rejected or revised. Therefore, the IOC result of the overall test was between 0.6 to 1 showing that every item was congruent with content validity and can be used to assess English proficiency.

- The 100 test items were then selected randomly based on the test specification by another 2 experts.

- The final 100 test items were trialed with 178 students to find the difficulty index and

reliability level. The overall difficulty is 0.246 with the reliability of 0.882.

After the pilot result, the 100 test items were then revised to adjust the items with high difficulty and then, trialed again. The outcomes of the main study were 0.95 reliability and 0.492 difficulty index, meaning a very high credibility and reliability.

#### 4. Results and Discussion

In this study, the findings are presented according to 3 research questions and 2 hypotheses.

**Research question 1: What are the concurrent validity values of the K-StEP Test scores and the TOEIC Test scores?**

In order to examine the relationship between the K-StEP Test scores and the TOEIC Test scores, the analysis of the correlation coefficient by means of Person's Product Moment was employed.

Hypothesis 1 was set to investigate the relationship between the K-StEP Test scores and the TOEIC Test scores.

Hypothesis 1: There is a significant relationship between the K-StEP Test scores and the TOEIC Test scores at the .05 level. ( $H_1 : r_{xy} \neq 0$ )

**Table 5** The Correlation Coefficient between the K-StEP Test scores and the TOEIC Test scores

		K-StEP	TOEIC
K-StEP	Pearson Correlation	1	.877**
	Sig. (2-tailed)		.00
	N	105	105

\*\* Correlation is significant at the .01 level (2-tailed).

\* Correlation is significant at the .05 level (2-tailed).

The findings from Table 5 showed that the correlation coefficient between the K-StEP Test scores and the TOEIC Test scores was .877 and it is statistically significant ( $p < .01$ ). A Pearson Product-Moment Correlation was run to determine the relationship between the K-StEP Test scores and

the TOEIC Test scores. There was a strong and positive correlation between the scores of the two tests ( $y = .877, N = 105, p < .01$ ).

It can be concluded that the K-StEP Test scores have a high positive correlation with the TOEIC Test scores. This means that if one variable

positively increases, there is a positive increase in another variable at the high level. Consequently, Hypothesis 1 stating that there is a significant relationship between the K-STEP Test scores and the TOEIC Test scores at the .05 level ( $H_1: r_{xy} \neq 0$ ) is accepted.

**Research question 2: To what extent can the K-STEP Test scores predict the TOEIC Test scores?**

Hypothesis 2 was established to examine if the K-STEP Test scores can predict the TOEIC Test scores.

Hypothesis 2: The K-STEP Test scores can significantly predict the TOEIC Test scores at the .05 level. ( $H_2 : B \neq 0$ ).

A Simple Regression Analysis was employed to assess the second hypothesis. Table 6 shows the model summary of the Simple Regression Analysis.

**Table 6** The Model summary of the Simple Regression Analysis

Model	R	R Square	Adjust R Square	Std. Error of the Estimate
1	.877 <sup>a</sup>	.770	.767	72.57

Predictors: (Constant), K-STEP scores  
Dependent variable: TOEIC scores

The R-coefficient of .877 was shown in Table 6. The R-Square was .770. This means that the independent variable can explain 77% of the

variance in the TOEIC Test scores. The standard error of the estimate is 72.57. Table 7 illustrates the coefficients of the regression model.

**Table 7** The Coefficients of the Regression Model

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-51.942	27.275	.877	-1.904	.060
	K-STEP	8.702	.469		18.552	.000

\*Dependent Variable: TOEIC scores

The unstandardized coefficient B is the increase or decrease in the dependent variable when an independent variable increases by one

unit. The regression equation used to predict the TOEIC Test scores can then be written as follows:

$$y = a + bx$$

$$\text{TOEIC Test scores} = -51.942 + 8.702 (\text{K-STEP Test scores})$$

According to the above equation, the K-StEP Test scores can be used to predict the TOEIC Test scores. With a unit increases in the K-StEP Test scores, the TOEIC Test scores will increase by 8.702 with a standard error of .469.

In summary, Hypothesis 2 which states that the independent variable can individually or in combination significantly predict the TOEIC scores at the .05 level ( $H_2: B \neq 0$ ) is accepted. If test takers know their K-StEP Test scores, it is possible that they will know the TOEIC Test scores.

**Research question 3: What are the ability bands for defining English levels obtained from the K-StEP scores?**

Since one of the objectives of the current study was to find the ability bands or descriptors for defining English proficiency obtained from the K-StEP Test scores and the TOEIC Test scores, the descriptors of CEFR levels of language proficiency were adopted in the study. In the present study, the descriptors of CEFR levels A1, A2, B1, and B2 were adopted to identify the English ability of the test takers at each level because they provide the band descriptors that can describe the performance and ability that the candidates or test takers of the K-StEP Test and the TOEIC Test are likely to be able to do.

By means of equation obtained, the ranges of the K-StEP Test scores that were converted to the TOEIC Test scores can be illustrated in Table 8.

**Table 8** The Ability Descriptors of the K-StEP Test scores equivalent to the TOEIC Test scores based on the descriptors of CEFR level A1, A2, B1 and B2

Number of Students	K-StEP	TOEIC	CEFR	Ability descriptors of CEFR
3 (2.8%)	20-31	120-224	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.
74 (70.47%)	32-68	225-549	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

Table 8 (Continued)

Number of Students	K-StEP	TOEIC	CEFR	Ability descriptors of CEFR
26 (24.76%)	69-95	550-785	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
2 (1.9%)	96 or above	786-944	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

Table 8 presents the K-StEP Test cut-off scores in relation to the TOEIC Test scores. The English proficiency of the students is interpreted based on CEFR ability bands which include “can do” statements. They describe what a foreign language learner can do at each level. In this study, it can be interpreted that there were only 2 students or 1.9% in B2 level. They achieved the K-StEP Test scores of over 96% and the TOEIC Test scores between 786 and 944.

There were 26 or 24.76% of the students in the B1 level, their ability having been described in Table 8. Most of the test takers, 74 students or 70.47% were in A2 level with the K-StEP Test scores between 32 to 68 and with the TOEIC Test scores ranging from 225 to 549. Though the test had been

designed to evaluate the English proficiency level A2, B1, and B2, there were 3 students or 2.8% who gained the K-StEP Test scores between 20 to 31 and the TOEIC Test scores from 120 to 224 or in A1 level.

In summary, once the K-StEP Test scores are reported, the test takers will know their TOEIC Test scores and their CEFR levels and understand the descriptors explaining what they can do in the ability descriptors of CEFR levels.

## 5. Discussion

### 5.1 The correlation between the K-StEP Test scores and the TOEIC Test scores

The high positive correlation between the K-StEP Test scores and the TOEIC Test scores means that if one variable positively increases, then a positive increase in another variable will occur. The

finding is in agreement with Anuyahongsa's (2015) study, which revealed that there was a very high positive correlation between the TNI English ability test scores and the TOEIC Test scores. Also, similar finding was also found in Jaturapitakkul and Watson Todd's (2018) report, which revealed that there was a high correlation between the TETET scores and the TOEIC Test scores. This means that the TETET and the TOEIC Test were very strongly correlated. In the current study, the structure of the K-STEP Test is quite similar to that of the TOEIC Test in that they include listening, reading, structure, and written expression with only 4-multiple choice items. Though the TOEIC test consists of a higher number of items-200 items, the K-STEP Test consists of only 100 items. However, the main purpose is to assess students' English proficiency. Another remarkable part is that the construction of the two tests is based on the CEFR descriptors that try to test what foreign language learners can do at each proficiency level. The study revealed highly positive correlation between the two tests which provided solid evidence that the K-STEP Test can effectively assess test takers' English proficiency in parallel to that of the TOEIC Test.

### 5.2 The equating predictive validity

The purpose of conducting a Simple Regression Analysis is to evaluate the relative impact of a predicted variable. This is different from a correlation analysis because it examines the degree of relationship between the two variables which in this study are the K-STEP Test scores and the TOEIC Test scores. According to the result of research question 2, it can be concluded that the K-STEP Test scores can be used to predict the TOEIC Test scores in the significant degree. The finding of the regression analysis was congruent with that of the concurrent validity. Similar finding was found in the study of Chanchusakun et al. (2007), which revealed that the SEPT scores can be used to predict the TOEIC Test scores at the significant level .05.

### 5.3 The ability bands to describe English proficiency levels of the K-STEP Test

After being mapped, the cut-off scores of the K-STEP Test were measured with respect to the CEFR levels and the TOEIC Test scores. The interpretation can be made, for example, as when a student achieves a score of 70 in the K-STEP Test, this student is considered to have an English proficiency equivalent to 600 in TOEIC or B1 of CEFR. Based on the CEFR global scale, this student can:

- Understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc.
- Deal with most situations likely to arise whilst travelling in an area where the language is spoken.
- Produce simple connected text on topics which are familiar or of personal interest.
- Describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans. (Council of Europe, 2011)

Stakeholders in KMUTNB students such as teachers, university administrators, or employers can use the information on cut-off score levels of the K-STEP Test in developing students' proficiency, designing curricula, setting targets or setting the entrance score or planning policy. The information can also be used to construct materials and plan classroom activities. Moreover, the students themselves will know their existing ability in English and can personally develop to meet the required level. Testing and evaluation is closely related to the process of learning and teaching. It aims not only evaluating the learners' proficiency, but also providing clear target for learners to develop themselves based on their competency. Therefore, standard testing is part of successful instruction to provide suitable learning levels and content.

## 6. Conclusion

This study has provided a summary of the development and validation of the K-StEP Test including the correlation between the K-StEP Test scores and the TOEIC Test scores. The findings revealed a highly positive correlation between the K-StEP Test scores and the TOEIC Test scores at .877. The results of the Simple Regression Analysis showed high equating validity between the two tests. Moreover, the results can be used to compare the test scores of the two tests and define the English ability of the test takers. The information on the cut-off score level can be used in developing students' proficiency and designing suitable courses for them.

### 6.1 Limitations of the study

Even though every effort was made to produce valid and reliable results of the study, there are several limitations that should be considered carefully.

- The first limitation is on the type of test.

The TOEIC Test measures general English language proficiency particularly in a working context while the K-StEP Test is designed for assessing English proficiency in academic endeavors. The number of items is also different. However, the 2 tests aim to assess test takers' language proficiency. Both TOEIC and K-StEP Tests can measure only receptive skills, i.e., listening and reading skills. Therefore, the test takers undertake the test based on only linguistic competence that is not similar to real performance. The interpretation of descriptors needs to be applied with caution.

- The second limitation is on CEFR descriptors.

Since CEFR descriptors are illustrative, not definitive which means that they can be interpreted differently by different experts. This means that different interpretation can lead to deviation in judgment of test items and cut-off scores. Thus, discussion among experts in the standard setting

process is necessary so that understanding and interpretation of the CEFR descriptors will be consistent.

### 6.2 Suggestions

Based on the above limitations, a future study using another method i.e. Angoff method, a group of panelists and other different standard tests could be conducted. Further research is needed to investigate the correlation and cut-off scores perhaps with other standard tests. The mapping process needs to be done carefully with higher numbers of experts. Training on the Angoff method is necessary before mapping. Also, the current cut-off scores should be considered by another panel of experts to increase credibility.

Another suggestion is that there should be a direct measure of writing skills and speaking skills in the K-StEP Test with which to gauge the real performance of the test takers according to the CEFR criteria.

## 7. Acknowledgements

This research was funded by the Thai Government Budget Expenditure, King Mongkut's University of Technology North Bangkok and the Faculty of Applied Arts. The authors would like to express their gratitude towards all participants and the Faculty staff for their kind cooperation and willingness to take part in this study.

## 8. References

- Anuyahongsa, B. (2015). Comparison of Efficiency between TNI English Ability Test and TOEIC Test. *Journal of Business Administration and Languages*. 3 (1), 1-6.
- Chanchusakun, S., Paiwithayasiritham, C., Polpanthin, Y. & Supannopaph, P. (2017). A Comparison of Quality of Equating English Scores between Equipercentile, Linear and Regression Methods. *Veridian E-Journal. Silpakorn University*, 10 (2), 2444-2455.



- Council of Europe. (2011). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*. Retrieved from [www.coe.int/lang-cefr](http://www.coe.int/lang-cefr)
- Education First (2018). *EF English Proficiency Index*. Retrieved from [www.ef.co.th/epi/](http://www.ef.co.th/epi/)
- Hirai, M. (2008). Correlation between STEP BULATS writing and TOEIC scores. Divergence and Convergence, *Educating with Integrity: Proceedings of the 7th Annual JALT Pan-SIG Conference*. Retrieved from <http://hosted.jalt.org/pansig/2008/HTML/Hirai1.htm>
- Hiranburana, H.Subphadoongchone, P. Tangkiengsirisin, S.Phoochaeoensil, S. Gainey, J. Thogsongsri , J. Sumonsriworakun, P. Somphong, M. Sappapan. P, & Taylor, P. (2017). A Framework of reference for English Language education in Thailand (FRELE-TH) - based on the CEFR, The Thai experience. *LEARN Journal*, 10 (2), 90-119.
- Jaturapitakkul, N. & Watson Todd, R. (2018). Test of English for Thai Engineers and Technologists (TETET) Technical Support. Retrieved from [www. Sola.kmutt.ac.th/ tetet/doc/TETET\\_Technical\\_report\\_NJ\\_RT\\_Complete.pdf](http://www.Sola.kmutt.ac.th/tetet/doc/TETET_Technical_report_NJ_RT_Complete.pdf)
- Martyniuk, W. (Ed). (2010). Aligning test with the CEFR: Case studies and reflections on the use of the Council of Europe's draft manual. *Studies in Language testing*. 33. Cambridge: Cambridge University Press.
- National Institute of Education Testing Services. (2018). *Manual for the 4th round external quality assessment in basic education*. Samutprakarn, Thailand: Offset Plus.
- Norris, J. M. (2006). The why (and how) of student learning outcomes assessment in college FL education. *The Modern Language Journal*, 90(4), 576-583. Retrieved from [https://doi.org/10.1111/j.1540-4781.2006.00466\\_2.x](https://doi.org/10.1111/j.1540-4781.2006.00466_2.x)
- North, B. (2014). *The CEFR in practice 4*. Cambridge: Cambridge University Press.
- Office of the Basic Education Commission. (2018). *Basic education core curriculum*. Bangkok.
- Tannenbaum, R.J. and Wylie, E.C. (2015). *Mapping TOEIC Test scores to the STANG language proficiency level*. Retrieved from [www.ets.org/toeic](http://www.ets.org/toeic)