



## Lesson Learned from the Initial Phase of the Development Process of CEFR-Based Uttaradit Rajabhat University Test of English Proficiency

Pittaya Yamo, Kunrawi Klinklan, Booranapong Kosenarak

Faculty of Humanities and Social Sciences, Uttaradit Rajabhat University

E-mail: [pittayachanai@yahoo.com](mailto:pittayachanai@yahoo.com), ORCID ID: <https://orcid.org/0000-0003-1594-8286>

E-mail: [kunrawi@yahoo.com](mailto:kunrawi@yahoo.com), ORCID ID: <https://orcid.org/0009-0009-1700-2823>

E-mail: [dom2544@yahoo.com](mailto:dom2544@yahoo.com), ORCID ID: <https://orcid.org/0009-0003-1078-3133>

Received 18/12/2023

Revised 20/12/2023

Accepted 25/12/2023

### Abstract

**Background and Aims:** The development of the CEFR-Based Test of English Proficiency enriches the broader field of language assessment. By sharing valuable insights into the complexities of aligning assessments with international frameworks. Thus, this academic article aims to present the lesson learned during the initial phase of the development process of the CEFR-Based Uttaradit Rajabhat University Test of English Proficiency (URU-TEP). Linking the URU-TEP to the Common European Framework of Reference for Languages (CEFR), widely used as a reference in English as a Foreign Language (EFL) education, is crucial if test results are to be interpreted under the CEFR levels.

**Methodology:** This study used the method of studying documents and related research, using thematic analysis and presenting descriptive information.

**Results:** the finding is divided into four sections: the current state of the English proficiency test at URU, the process of developing the five-skill URU-TEP based on the CEFR, the process of obtaining test validity and reliability, and opportunities and challenges.

**Conclusion:** The essence of the discovery is to enhance understanding in various areas. of the English proficiency testing process at Uttaradit Rajabhat University It also emphasizes consideration of the validity and reliability of the test to support quality and improve the examination process in the future.

**Keywords:** Test of English Proficiency; Test Validity; Test Reliability; CEFR-based Test

## บทเรียนที่ได้รับจากขั้นตอนแรกของการพัฒนาข้อสอบวัดความสามารถภาษาอังกฤษของมหาวิทยาลัย ราชภัฏอุตรดิตถ์ตามกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป

พิทยา ยาไม้, กุสรวี กลิ่นกลิ่น, บุรณพงษ์ โกเสนารักษ์

คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏอุตรดิตถ์

### บทคัดย่อ

**ภูมิหลังและวัตถุประสงค์:** การพัฒนาแบบทดสอบความสามารถทางภาษาอังกฤษตาม CEFR ช่วยเสริมการประเมินภาษาในสาขาที่กว้างขึ้น ด้วยการแบ่งปันข้อมูลเชิงลึกที่มีคุณค่าเกี่ยวกับความซับซ้อนของการประเมินที่สอดคล้องกับกรอบการทำงานระหว่างประเทศ ดังนั้นบทความวิชาการนี้มีวัตถุประสงค์เพื่อนำเสนอบทเรียนที่ได้รับจากขั้นตอนแรกของการพัฒนาข้อสอบวัดความสามารถภาษาอังกฤษของมหาวิทยาลัยราชภัฏอุตรดิตถ์ตามกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป ซึ่งการอ้างอิงกรอบอ้างอิงความสามารถดังกล่าวมีการใช้กันอย่างแพร่หลายในการศึกษาภาษาอังกฤษเป็นภาษาต่างประเทศ โดยเฉพาะหากต้องมีการตีความผลการทดสอบจากข้อสอบที่พัฒนาขึ้นตามกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป

**ระเบียบวิธีการศึกษา:** การศึกษาค้นคว้าครั้งนี้ใช้วิธีการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ทำการวิเคราะห์แก่นสาระแล้วนำเสนอเชิงพรรณนาความ

**ผลการศึกษา:** บทความวิชาการนี้แบ่งออกเป็นสี่ส่วน ได้แก่ สถานะปัจจุบันของการทดสอบความสามารถทางภาษาอังกฤษ ณ มหาวิทยาลัยราชภัฏอุตรดิตถ์ กระบวนการพัฒนาข้อสอบทั้งห้าทักษะตามกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป กระบวนการตรวจสอบความเที่ยงตรงและความน่าเชื่อถือของแบบทดสอบ และโอกาสและความท้าทาย

**สรุปผล:** สาระสำคัญของการค้นพบเป็นการเสริมสร้างความเข้าใจในด้านต่างๆ ของกระบวนการทดสอบความสามารถทางภาษาอังกฤษ ณ มหาวิทยาลัยราชภัฏอุตรดิตถ์ อีกทั้งยังเน้นการพิจารณาความเที่ยงตรงและความน่าเชื่อถือของแบบทดสอบเพื่อสนับสนุนคุณภาพและปรับปรุงกระบวนการตรวจสอบในอนาคต

**คำสำคัญ:** ข้อสอบวัดความสามารถภาษาอังกฤษ; การตรวจสอบความเที่ยงตรง; การตรวจสอบความเชื่อมั่น; ข้อสอบตามกรอบอ้างอิงความสามารถในการเรียนภาษาของสภายุโรป

### Introduction

#### Current State of English Proficiency Test at Uttaradit Rajabhat University

In 2019, Uttaradit Rajabhat University (URU) agreed to set an English proficiency standard for bachelor's degree level students, by the Office of the Higher Education Commission's announcement on the policy to raise the standard of English in higher education institutions dated 19 April 2016. The policy's objective is to promote the English proficiency of students in all university programs so that they are ready for academic, professional, and English

communication skills that are usable and comparable to international standards. URU then requires graduates to have a CEFR-equivalent score on any standardized test of English proficiency. This means that URU graduates with an English major must have a B2, whereas non-English majors must have a B1.

Because URU does not yet have a standardized English proficiency test, the university is now using a variety of English test providers and tests. They include English Discoveries, British Council English Scores, TOEIC (Test of English for International Communication), TOEFL-ITP (Test of English as Foreign Language - Institutional Testing Program), CU-TEP (Chulalongkorn University Test of English Proficiency), TEC-W (Test of English Communication in the Workplace) and so on. English Discoveries is primarily used as a provider of English proficiency tests with a monthly test schedule in collaboration with the URU Language Center, and test takers are responsible for the test fee. However, graduates may still use the CEFR-B1 test scores from currently available standardized tests to prove their English language proficiency to satisfy the graduation requirement. Below are the cut-off scores for standardized tests that correspond to the CEFR-B1 level (Wudthayagorn, 2018). If a student at URU receives a test score in any of the following ranges, they will be able to graduate and satisfy the university's English proficiency requirement.

**Table 1** The Standardized Test Cut-off Scores equivalent to the CEFR-B1 Level

CEFR Level	TOEFL iBT	TOEFL ITP	IELTS	TOEIC	TOEIC Bridge	CU-TEP
B1	57	460	5.0	790	170	35

According to what was presented above, the university seems to expect lecturers of English to develop the URU-TEP as a standardized test under the close supervision of an expert. The objective of this article is to discuss the initial phase of the CEFR-Based Uttaradit Rajabhat University Test of English Proficiency (URU-TEP) development process. It is divided into four sections: opportunities and challenges, the five-skill URU-TEP development process based on the CEFR, the current state of the English proficiency test at URU, and the process of obtaining test reliability and validity. The next section discusses the process of developing the four-skill URU-TEP based on the CEFR.

## Development Process of URU-TEP based on Common European Framework of Reference for Languages

In EFL English education, including Thailand, the Common European Framework of Reference for Languages, or CEFR, is frequently used as a reference. It is an international language ability standard. It consists of assigning a six-point scale to language ability, ranging from A1 for beginners to C2 for those who have fully mastered a language. This allows anyone involved in language teaching and testing, such as teachers or students, to quickly determine the level of various qualifications (CambridgeUniversityPress&Assessment, 2023). As shown in the table below, the Council of Europe (2023) summarized the six-point scale of language ability from A1, basic user, to C2, proficient user.

**Table 2** Common Reference Levels: Global Scale (Council of Europe, 2023)

Level Name	CEFR Level	Can-do Statements
Proficient User	C2	<ul style="list-style-type: none"> <li>● Can understand with ease virtually everything heard or read.</li> <li>● Can summarise information from different spoken and written sources, reconstructing arguments, and accounts in a coherent presentation.</li> </ul> <p>Can express him/herself spontaneously, very fluently, and precisely, differentiating finer shades of meaning even in more complex situations.</p>
	C1	<ul style="list-style-type: none"> <li>● Can understand a wide range of demanding, longer texts, and recognize implicit meaning.</li> <li>● Can express him/herself fluently and spontaneously without much obvious searching for expressions.</li> <li>● Can use language flexibly and effectively for social, academic, and professional purposes.</li> </ul> <p>Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors, and cohesive devices.</p>

Level Name	CEFR Level	Can-do Statements
Independent User	B2	<ul style="list-style-type: none"> <li>● Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization.</li> <li>● Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party.</li> <li>● 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.</li> </ul>
	B1	<ul style="list-style-type: none"> <li>● Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc.</li> <li>● Can deal with most situations likely to arise whilst traveling in an area where the language is spoken.</li> <li>● Can produce simple connected text on topics that are familiar or of personal interest.</li> <li>● Can describe experiences and events, dreams, hopes &amp; ambitions and briefly give reasons and explanations for opinions and plans.</li> </ul>
Basic User	A2	<ul style="list-style-type: none"> <li>● 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).</li> <li>● Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters.</li> <li>● Can describe in simple terms aspects of his/her background, immediate environment, and matters in areas of immediate need.</li> </ul>



Level Name	CEFR Level	Can-do Statements
	A1	<ul style="list-style-type: none"><li>● Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type.</li><li>● 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.</li><li>● Can interact simply provided the other person talks slowly and clearly and is prepared to help.</li></ul>

In March 2023, a project funded by Uttaradit Rajabhat University began to develop the URU-TEP by the Common European Framework of Reference for Languages (CEFR). The 80-question test was developed with the CEFR-B1 level descriptors as a guideline, which is shown in the table above. URU lecturers of English from the Education Faculty, the Humanities and Social Sciences Faculty, and the International College were invited to participate in this project to develop the five parts of the test: listening comprehension, speaking, vocabulary, writing & grammar, and reading comprehension based on their expertise and interests. Details on each part are provided as follows.

The first part, Listening Comprehension, consists of 15 questions (Questions 1-15). It consists of five listening items: two announcements and three conversations, each of which the test taker will be asked three questions. The test taker will be asked the questions with four choices. After hearing each announcement and conversation, they must answer three questions in 30 seconds by selecting the best option based on the questions. The questions and choices will be on their answer sheet, and each announcement and conversation will be repeated twice.

The second part, Speaking, is an indirect test that assesses a test taker's ability to communicate. There are a total of 10 questions (Questions 16-25). This part includes five Question-Response pairs and three dialogues or conversations.

The third part, Vocabulary, contains 10 questions (Questions 26-35). This part requires a test taker to locate the appropriate vocabulary and synonyms for parts of speech to complete the sentences.





The fourth part, Writing and Grammar, contains 25 questions (Questions 36-60). This part requires a test taker to choose the grammatically incorrect part of each sentence from the four underlined parts.

The fifth and last part, Reading Comprehension, consists of 20 questions (Questions 61-80). The test taker will read three short texts and one long text. Each short text is followed by four questions, and each long text by eight questions.

A standard-setting method must be taken into account to verify the validity and confirm the reliability of the above URU-TEP test questions. The process for obtaining test validity and reliability is addressed in the following section. To create the test questions for each of the mentioned parts, a group of four URU English language lecturers, known as test writers, were free to form as they saw appropriate and were based on their shared interests and areas of expertise. To choose the appropriate online texts that were considered authentic texts, a group member needed to comprehend the CEFR-B1 descriptors in detail. After completing a first draft of test question development, each group was required to present their questions to the external expert who had been invited to participate in this project as well as other group members, known as member checking, for feedback and suggestions for a better version. The test writers must remain the same as suggested by the external expert. This means that those responsible for a specific part should continuously work together because they are likely to be familiar with their test questions in terms of how to create the test and where to find authentic texts if there are multiple sets of test questions. This ensured that the questions were ready for the next phase of determining test validity and reliability.

### Process of Obtaining Test Validity and Reliability

The constructed URU-TEP test questions will be determined by the 6-10 experts/judges using the Angoff method, and they will be piloted with URU students to ensure test validity and reliability, respectively. The test must be designed to measure what it claims to measure through the YES/No Angoff method to be considered for CEFR-B1 test validity. To connect the high-stakes standardized tests with the CEFR, the YES/NO Angoff is frequently used in standard settings (Nakanitanon, 2021). Numerous standardized tests, including the TOEFL iBT (Tannenbaum & Wylie, 2008), TOEFL ITP (Tannenbaum & Baron, 2011), TOEFL Junior Comprehensive Test (Tannenbaum & Baron, 2015), and TOEIC (Tannenbaum & Wylie, 2019), use this method. The fundamental steps of the Yes/No Angoff method, according to





Nakanitanon (2021), are the selection of qualified judges, the experts/judges' evaluation of the borderline test-takers, averaging the judgment scores, analysis of the data for the mean cut score, discussion of the mean cut score, and agreement on the cut score attained. This method appears to be similar to the IOC approach in that experts/judges only provide yes/no estimates, scoring "Yes" as one point and "No" as zero. Six experts/judges will be invited to determine test validity for the development of the URU-TEP test. According to Sridhanyarat et al. (2021), the following were the responsibilities of the experts/judges. First, they needed to determine whether the test questions corresponded to the test specification's objectives. Then they had to see if the test items were grammatically correct and natural to native English speakers. It should be noted that all steps were taken by the proposed method. It was indexed at 0.75, which helped ensure that a specific test item was content-valid.

In terms of test reliability, it refers to the consistency of a test-taker's scores on the same test taken twice. Test-retest results indicate temporal stability and split-half results show the test's internal consistency. This means that giving test takers a pilot test once is insufficient for test reliability. URU students will take part in the pilot phase of the development of the URU-TEP test. A statistician will examine the test questions after they are administered to determine the index of difficulty and index of discrimination. According to Sridhanyarat et al. (2021), to determine whether a particular item is actually at the level of difficulty anticipated, piloting with suitable samples of test takers is imperative. Additionally, it enables us to gather as much feedback as possible on the test instrument's quality. This means that piloting is a technique for testing the test instrument to clear up any ambiguities and assess the readability and clarity of test items. The task difficulty level was interpreted during this process, and time overload was assessed. However, because the URU-TEP test is still being developed or is just beginning, a conclusion with a statistical presentation for test validity and reliability has not yet been finalized.

## Opportunities and Challenges

While developing a standardized test that employs the CEFR, like the URU-TEP test, opportunities and challenges might coexist. Specifically, how lecturers assist their students in developing their English in preparation for the test. In terms of opportunities, we can envision what kinds of teaching and learning activities are required to begin once a standardized test is ready for test-takers. This is viewed as an opportunity for both instructors and students to







learn how to teach and improve a specific set of skills that are connected to the test. The "can-do" descriptors given by the Council of Europe (2023) are probably a good place to start based on the CEFR-B1. The four "can-do" statements, (1) can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc., (2) can deal with most situations likely to arise whilst traveling in an area where the language is spoken, (3) can produce simple connected text on topics which are familiar or of personal interest, and (4) can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans, are considered an intermediate level. The topics covered at this CEFR level are ones that students will use daily, such as at work, school, and during free time.

Lecturers appear to revisit their lessons by taking the four CEFR-B1 "can do" statements mentioned above into consideration. In other words, they need to create appropriate course material, redesign the tasks and assignments, and determine appropriate course evaluation, and they should be interrelated. It is evidenced by Hawes (2020) that "CEFR offers an opportunity to look at planning, teaching, and assessment from different perspectives and see how they all link up" (p. 75). In response to this, URU English language lecturers, for example, agreed to use a series of commercial textbooks for teaching general education courses 1-3 because these textbooks assist students in developing their English skills in a sequence-based manner, from CEFR-A1 to CEFR-B1. They then appear to have a chance to grow personally by using high-quality, conventional textbooks that are in line with the CEFR. However, in a different environment such as Thailand, using commercial textbooks produced in the West might not fit a student's learning style. Particularly, some parts of the commercial ones might not help students better understand what is necessary to, for instance, obtain satisfactory results from the CEFR-B1 test. This means that tailoring individual textbooks to the needs and preferences of the students using a variety of books or authentic online sources could assist instructors in finding a solution (Hawes, 2020).

Lecturers have an opportunity to take into account the four CEFR-B1 "can do" statements when developing the URU-TEP test by the CEFR. For example, when creating the test questions for the Vocabulary Part using authentic sentences made available online, they must keep the CEFR-B1 vocabulary list in mind. They might gain a better understanding of the CEFR and maintain their progress toward the CEFR-B1 level with the support of this practice. One more illustration is for the reading part. Lecturers are required to select a short article from an online





source, to map it to the B1 level descriptors. The selected text must then be evaluated by a CEFR readability checker, such as one found at <https://readabilityformulas.com>, which counts the words, sentences, syllables, and characters to determine whether it is appropriate for the CEFR-B1 reading level. Overall, aligning the URU-TEP test with the CEFR could provide URU lecturers an opportunity to gain a deeper understanding of what the CEFR is at each level. They can create textbooks that are in line with the desired CEFR level or choose the appropriate online text.

Due to the challenges in creating the URU-TEP test by the CEFR, lecturers may become unmotivated by the huge amount of work they must perform. It is a difficult task to develop such a standardized test from scratch. Thus, what is viewed as opportunities above may put stress and pressure on those involved. Since the development of a standardized test involves many phases and requires time and effort, those involved in the project may experience stress and anxiety because it adds to their already heavy workload. The challenges perceived by instructors were derived from many factors, such as inadequate training to familiarize instructors with the CEFR system (Hawes, 2020), support to meet a language policy (Mohamad Uri, 2023), expectations of students' English proficiency (Sridhanyarat, Pathong, Suranakkharin, & Ammaralikit, 2021), and expected outcomes of the policy to increase English ability (Wataksorn, 2021). In addition to facing these challenges, URU instructors' motivation is likely to be increased if they are given incentives and support to take part in the project.

## Conclusion

In conclusion, developing a standardized test based on the CEFR entails multiple phases, as well as time and effort. This article only covered the first phase by first showing how the English proficiency test is administered to students who need to pass it to graduate at Uttaradit Rajabhat University. It then highlighted how to determine the test validity and reliability, providing an overview of the process of developing the test questions for the five parts. To ensure its effectiveness, this phase must be repeated several times. The opportunities and difficulties were then discussed. Based on what was presented above, an ideal flowchart of the URU-TEP test development was then proposed below.



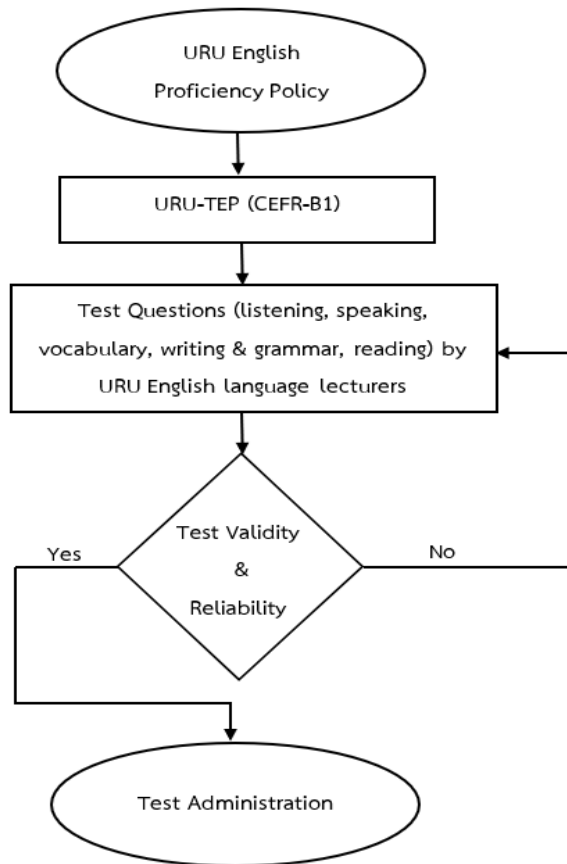


Figure 1 Proposed URU-TEP Test Development

## References

- CambridgeUniversityPress&Assessment. (2023). *English language assessment*. Retrieved from <https://www.cambridgeenglish.org/exams-and-tests/cefr/>
- CouncilOfEurope. (2023). *Global scale: Common reference levels*. Retrieved from <https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale>
- Hawes, C. (2020). CEFR & language teaching standardization at Suratthani Rajabhat University. *Journal of Humanities and Social Sciences*, 12(2), 59-77.
- Mohamad Uri, N. (2023). Challenges in CEFR adoption: Teachers' understanding and classroom practice. *International Journal of Modern Languages and Applied Linguistics*, 7(1), 49-62.
- Nakanitanon, P. (2021). Linking an English proficiency test to the CEFR: Setting valid cut scores. *The New English Teacher*, 15(1), 53-64.



- Sridhanyarat, K., Pathong, S., Suranakkharin, T., & Ammaralikit, A. (2021). The development of STEP, the CEFR-Based English proficiency test. *English Language Teaching*, 14(7), 95-106.
- Tannenbaum, R., & Baron, P. (2011). *Mapping TOEFL ITP scores onto the Common European Framework of Reference*. Princeton: NJ: Educational Testing Service.
- Tannenbaum, R., & Baron, P. (2015). *Mapping scores from the TOEFL Junior Comprehensive test onto the Common European Framework of Reference (CEFR)*. Princeton: NJ: Educational Testing Service.
- Tannenbaum, R., & Wylie, E. (2008). *Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology*. Princeton: NJ: Education Testing Service.
- Tannenbaum, R., & Wylie, E. (2019). *Mapping the TOEIC Tests on the Common European Framework of Reference*. Princeton: NJ: Education Testing Service.
- Wudthayagorn, J. (2018). Mapping the CU-TEP to the Common European Framework of Reference (CEFR). *LEARN Journal: Language Education and Acquisition Research Network*, 11(2), 163-180.

