

Aligning the CEFR Levels with the KU-EPT Scores

Sun-Young Shin
Panjanit Chaipuapae
Chutima Bunparit

July 27, 2024

Kasetsart University Language Testing Center
Faculty of Humanities Kasetsart University

Aligning the CEFR Levels with the KU-EPT Scores

Principal Investigator (PI): Sun-Young Shin¹

Co-PIs: Panjanit Chaipupae and Chutima Bunparit²

¹ Indiana University ² Kasetsart University

The Kasetsart University English Proficiency Test (KU-EPT) is a standardized test designed to assess general academic English proficiency. It is intended for Kasetsart University graduate students as well as for members of the public who wish to evaluate their English skills for personal development, job applications, career advancement, or further education opportunities. The KU-EPT consists of three main parts: Structure (25 items), Cloze test (15 items), and Reading (40 items). All the test items in the KU-EPT are based on the multiple-choice format with four options. The entire KU-EPT has a score range of 0 to 80, which is converted into a percentage.

This report presents the findings of an empirically-based linking study aimed at aligning KU-EPT scores with the widely recognized Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001, 2020) levels. The study utilized the Yes/No Angoff method (Fechter & Yoon, 2024; Hsieh, 2013; Impara & Plake, 1997), renowned for its extensive use and thorough research in standard settings across various contexts (Cizek & Bunch, 2007). This method is favored for its ability to generate multiple cut scores efficiently and with lower cognitive demands on panelists yet yielding comparable results to other standard-setting methods (Plake & Cizek, 2012).

The study aimed to establish four distinct cut scores corresponding to the four CEFR levels (A2, B1, B2, and C1). Due to the limited number of test items and the scarcity of test takers at the A1 and C2 levels, differentiation between A1 and A2 as well as C1 and C2 was not feasible. The Principal Investigator (PI) led and oversaw the entire project, with the Co-PIs responsible for the preparation of KU-EPT data and materials. The Co-PIs coordinated a panel of 10 experts familiar with the KU-EPT and the target test-taker population. Currently, there are 13 different versions of the KU-EPT available, among which one version was selected for standard-setting purposes based

on its highest internal consistency reliability value ($\alpha = .900$), a large number of test takers ($N = 302$), and its average level of difficulty ($M = 38.66$, $SD = 12.50$).

This standard-setting study was conducted both online via Webex and in person. The study was organized into three parts: Session One was completely online; in Session Two, only the PI was online while the panelists convened onsite; and Session Three was completely onsite (see Appendix). A total of 10 panelists, with extensive English teaching experience and familiarity with the KU-EPT and the target population, were recruited and divided into two subgroups of five panelists each with the Co-PIs as a team leader of each group, following guidelines suggested by the standard-setting literature (Loomis, 2012; Shin & Lidster, 2017; Tannenbaum & Cho, 2014;). Table 1 provides demographic information on the 10 panelists.

Table 1

Demographic Information on the 10 Panelists

Demographics		Number (percentage)
Gender	Male	1 (10%)
	Female	9 (90%)
Academic Title	Assistant Professor	3 (30%)
	Associate Professor	3 (30%)
	None	4 (40%)
Education	Master's degree	1 (10%)
	Doctoral degree	9 (90%)
Field of study	Linguistics, Translation studies, English Language Studies, Teaching English as a Foreign Language, Languages, Applied Linguistics	10 (100%)
Teaching experience	6-10 years	1 (10%)
	11-15 years	3 (30%)
	16-20 years	2 (20%)
	Over 20 years	4 (40%)

During the initial Webex meeting, each panelist received documents containing abridged CEFR-level descriptors for each language skill. The PI then introduced these descriptors to ensure a common understanding of English language proficiency benchmarks among the panelists. Additionally, he provided detailed information about the KU-EPT, including sample passages and items, to ensure that the panelists were familiar with and shared the common understandings of the test content and the CEFR-level descriptions. Subsequently, a practice session was held to prepare the panelists for the yes/no Angoff standard-setting method, during which they determined cut scores for each CEFR level using sample KU-EPT items.

KU-EPT Cut Scores

Between the online and in-person meetings, each panelist generated the three cut-offs to be mapped onto the CEFR levels using the Yes/No Angoff in the first round. In the Yes/No Angoff method, the panelists conceptualized a borderline test taker who possesses a minimum level of the four CEFR levels (A2, B1, B2, and C1) and then judged whether the borderline test takers at each level would answer each item correctly (Yes equal to 1) or incorrectly (No equal to 0). In the first round of the Yes/No Angoff standard-setting procedure, the panelists completed their individual yes/no judgments for all items. During the in-person meeting, each panelist's individual three-cut scores were compared with others, and they received feedback on their judgments. As part of the feedback and panel discussion within each subgroup, information about other panelists' cut scores and the difficulty of test items for the takers was provided to guide their cut score setting in the second round. Following these discussions, the panelists resumed their ratings for each item in the second round. After completing the second round of ratings, the PI assessed the levels of agreement on each cut score generated by the two subgroups. Subsequently, he discussed the second-round ratings with the Co-PIs to finalize the cut-offs for each CEFR level. Table 2 below presents the cut scores for each of the four CEFR levels generated using the Yes/No Angoff method. The CEFR descriptors for KU-EPT in terms of grammatical accuracy, vocabulary range, and overall reading comprehension are summarized in Table 3.

Table 2*Total KU-EPT Cut Scores for CEFR levels*

CEFR levels	KU-EPT Raw Scores (Percentage)
A2	5 – 32 (6 – 40%)
B1	33 – 63 (41 – 79%)
B2	64 – 72 (80 – 90%)
C1	73 – 80 (91 – 100%)

Table 3*CEFR Descriptors for KU-EPT*

Proficient user	C1	<p>Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.</p> <p>Can understand and use appropriately the range of technical vocabulary and idiomatic expressions common to their area of specialization.</p> <p>Can understand in detail lengthy, complex texts, whether or not these relate to their own area of speciality, provided they can reread difficult sections.</p> <p>Can understand a wide variety of texts including literary writings, newspaper, or magazine articles, and specialised academic or professional publications, provided there are opportunities for rereading, and they have access to reference tools.</p>
Independent user	B2	<p>Has good grammatical control; occasional “slips” or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.</p> <p>Shows a relatively high degree of grammatical control.</p> <p>Has a good command of simple language structures and some complex grammatical forms.</p> <p>Has a good range of vocabulary for matters connected to their field and most general topics.</p> <p>Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary but may experience some difficulty with low-frequency idioms.</p>
		B1

Basic User	A2	<p>Uses some simple structures correctly, but still systematically makes basic mistakes.</p> <p>Has sufficient vocabulary to conduct routine everyday transactions involving familiar situations and topics.</p> <p>Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency every day or job-related language.</p> <p>Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.</p>
-------------------	-----------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

References

- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment. Companion volume*. Strasbourg: Council of Europe. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>.
- Fechter, T. M., & Yoon, H. (2024). Evaluating methodological enhancements to the Yes/No Angoff standard-setting method in language proficiency assessment. *Language Testing, onlinefirst*.
- Hsieh, M. (2013). Comparing Yes/No Angoff and Bookmark standard setting methods in the context of English assessment. *Language Assessment Quarterly, 10*(3), 331-350.
- Impara, J. C., & Plake, B. S. (1997). Standard setting: An alternative approach. *Journal of Educational Measurement, 34*, 353-366.
- Loomis, S. C. (2012). Selecting and training standard setting participants: State of the art policies and procedures. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 107-134). New York, NY: Routledge.
- Plake, B. S., & Cizek, G. J. (2012). Variations on a theme: The modified Angoff, extended Angoff, and Yes/No standard setting methods. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 181-199). New York, NY: Routledge.

Shin, S.-Y., & Lidster, R. (2017). Evaluating standard setting methods in an ESL placement testing context. *Language Testing*, 34(3), 357-381.

Tannenbaum, R. J., & Cho, Y. (2014). Critical factors to consider in evaluating standard setting studies to map language test scores to frameworks of language proficiency. *Language Assessment Quarterly*, 11(3), 233-249.

