Developing a Blueprint for English Proficiency Test with Vocational Nuance Based on CEFR

Original Article

Gullit Tornado Taufan[1*], Rizqi Febrian Pramudita[2], Meiga Rahmanita[3], Nur Afni Rachman[4], Astrid Hidayanti Ananda Putri[5]

[1-5]Language, Communication, and Tourism Department, Politeknik Negeri Jember, Jember, Indonesia
Email: [1] gullittaufan@polije.ac.id, [2] rizqi.febrian@polije.ac.id, [3] meiga.rahmanita@polije.ac.id, [4] nur.afni@polije.ac.id, [5] astridhidayanti_ap@polije.ac.id

Abstract

Standardized English proficiency test has been an essential part in English language teaching. Despite the fact that globally recognized tests such as TOEFL, IELTS, and TOEIC are, still, dominating, a number of institutions in Indonesia have started to administer their English proficiency tests. One of which is Politeknik Negeri Jember (Polije), a higher educational that focuses on vocational education. In line with its characteristics, developing standardized English proficiency test with vocational nuance should be more acceptable. Therefore, the development of the blueprint of English proficiency test with vocational nuance is conducted through this research and development (R&D). In this case, the blueprint is based on the Common European Framework of Reference as the basis of the development. The initial stage of the development, which focused on the analysis of important aspects of the development of the blueprint, has been completed. The stage is followed by the development stage, which is mainly discussed in this report, by considering the development aspects necessary for the new blueprint. As a result, the new blueprint of EPT Polije is developed and ready to be implemented.

Keywords: Blueprint, CEFR, Testing, Proficiency, Vocational.

1. Introduction

English proficiency test is undoubtedly an essential part of English language teaching (ELT). Such tests are designed to measure the test takers' English proficiency covering the four skills of English. In Indonesia context, where English is spoken as a foreign language, TOEFL and IELTS are considered to be the globally recognized, especially in academic settings (Roza, 2019), while TOEIC is appropriate for professional and workplace contexts. In most cases, the purposes of taking the tests are to fulfill administrative requirements including university or college admission, scholarship qualification, professional certification, job application (Susyla et al., 2021), etc. Therefore, it is undeniably acceptable that the standardized tests are still relevant to certain purposes and needs.

Despite the fact that standardized English proficiency test is necessary in certain contexts and purposes, there are issues surrounding the, so-called, popular standardized tests. One of the questions arising is regarding the relevance of English standardized tests with the backgrounds of the test takers including the cultural aspects that are not relevant to the them and content focus which seems to be more Western-centered (Ojochegbe, 2024). Similarly, Yasmin (2024) explains that the popular standardized tests contain of Western-centric themes which are less relevant to, specifically mentioned, Asians. Those, to certain extend, happens to most Indonesians where English is spoken as a foreign language. The next issue to consider

is the affordability of taking the tests. Several studies suggest that the internationally-recognized standardized tests are far from being affordable for many test takers (Arnelli, 2025; Azhari et al., 2021; Yasmin, 2024) including Indonesians. In that case, the costs do not only cover the official fees of the tests but the preparations as well. Test takers should be well-prepared prior to the tests to avoid failure; they might have to pay double to take the same tests in the future.

It is clear that the discussion about related issues regarding the English standardized tests is still relevant in general contexts. Nevertheless, empirical studies regarding the administration of English standardized tests in vocational contexts should also be considered essential as the standardized tests are expected to reflect both English competence and vocational relevance. By focusing specifically on this issue, including the process of developing a blueprint of English proficiency test, the gap of the less-explored area can be answered. In this case, the structured investigation of the development as well as the necessary process should be presented to achieve a clear scientific information.

Based on the similar issues as explained above, Politeknik Negeri Jember (Polije) has started to administer its institutional English proficiency test (EPT) under the Language Center Polije. The EPT Polije has run for several years, based on the interview with the head of Language Center Polije, and continues development is very much necessary. Therefore, to investigate what development needed at the time being, an initial stage of this study has been completed. Through interview and document analysis, it was found that fundamental adjustment is crucial. In this case a new blueprint of EPT Polije, as the guideline for developing the test, was proposed.

The development of the new blueprint of EPT Polije, as a result of the initial stage of the study, covers the format of the blueprint, number of questions or items, vocational aspects to be included in the test, and distribution of levels of difficulties based on CEFR. Based on the results of the analyses in the previous stage, the development of the new blueprint could be executed. Therefore, this report focuses on the process of the development of the new blueprint of EPT Polije, by briefly reviewing the initial stage as an unseparated part of the overall process.

2. Literature Review

2.1. Standardized English Proficiency Test

English proficiency test has been recognized for a long moment of time as important tool or method to measure test takers' knowledge and skills in English both in academic and professional or occupational contexts. Globally, TOEFL and IELTS are the most well-known standardized tests to evaluate listening, speaking, reading, and writing skills (Ojochegbe, 2024) for various purposes such as college admission, graduation, scholarship, etc. (Roza, 2019). With the decades of experience in administering English standardized tests, those are believed to demonstrate strong validity and reliability. Ockey and Gokturk (2019) state that such standardized tests are developed based on careful and scientific investigation to achieve the empirical support.

Regardless of the importance of internationally recognized standardized English proficiency tests in measuring ones' English proficiency, there are critiques surrounding their existence. For instance, Ockey and Gokturk (2019) argue that their claims over test takers' abilities, as suggested by their scores, may not always picture their true English proficiencies. They further explain that various factors including their experience in taking such tests and their backgrounds that may vary across different disciplines and, even, institutions. The next

issue is related to the cultural and linguistic diversity of the test takers which are considered unfair for English non-native speakers who come from non-Western countries (Ojochegbe, 2024). Besides, it is also explained in the previous part that taking such standardized tests may cost a lot for the test takers, not only for the official prices of the tests but also for the preparation (Arnelli, 2025; Azhari et al., 2021; Yasmin, 2024).

Concerning the issues explained above, the implementation of so-called popular standardized tests may not be, what Bachman and Palmer (2010) explain, inappropriate for the test takers especially when they are exposed to English in a specific situation. For instance, the standardized tests which is known best for academic settings may not be appropriate to be implemented for test takers who study at a vocational institutions. Therefore, designing a test that can accommodate test takers' specific characteristics would provide them more appropriate and fair assessment. In this case, in line with the focus of this study, one of the relevant issue to accommodate is developing English proficiency test in vocational setting.

2.2. Blueprint of English Proficiency Test

As implied in the previous section, the development of standardized English proficiency test requires serious process to achieve the expected validity and reliability. In this case, the process should start from the very initial stages prior to the construction of the test items. In relation to the focus of this study, one of the initial stages is developing the blueprint of the test. Fulcher (2010) explains that the blueprint, which is also called test specification, provides the detailed architecture of a test to show how to construct it. Further, he explains that one of the most important aspects is that the blueprint should provide clear clues on what abilities or proficiencies are tested.

Developing blueprint of a standardized English proficiency test is crucial. As noted by Cheewasukthaworn (2022), this is highly significant in the development process of a test for this provides purpose and specifications of the items. Another important aspect of the blueprint is that this provides a test developer a reference to control test tasks, from which when new versions, but equivalent, of the test are needed, they can be developed, trialed, and normed (Davidson, 2012). It is obvious that the blueprint functions as the guidelines of the development of, so-called, controlled English proficiency test.

With regard to this study, it is also important to note that the blueprint being developed is related to a specific setting. The blueprint of EPT Polije is intended to be implemented in constructing English proficiency test in vocational context. This is in line with Bachman and Palmer's (2010) explanation on correspondence between language use and language test performance. They explain that there is correlation between language test performance and the characteristics of the test takers. In this case, the blueprint of EPT Polije targets, primarily, the students of the vocational institution.

2.3. Common European Framework of Reference (CEFR)

Common European Framework of Reference (CEFR) has been an important reference in English language teaching and assessment. Developed by Council of Europe in mid-1990s, the document has provided stakeholders in language education with a foundation to help them pursue certain objectives; one of which is developing a standardized English proficiency test (Piccardo, 2019). Being a transparent, coherent, and comprehensive basis for, one of which, developing language testing, the CEFR describes English proficiency in six levels: A1, A2, B1, B2, C1, and C2 (Council of Europe, 2025) in which 'A' is for basic user, 'B' for independent user, and 'C' for proficient user (Cheewasukthaworn, 2022). By relating to CEFR's descriptors, it is possible for a test developer to design a test with intended levels of proficiency, which is related to the levels of difficulty. In this case, specifically focusing on the blueprint being

developed, the CEFR is the main reference regarding the proficiency levels as well as distribution of the test items.

3. Methods

As implied in the previous sections, this study is a research and development (R&D) with the focus of developing the blueprint of EPT Polije. The 'research' part, which is also called initial stage, of this study has been done and reported while this focuses mainly on the development of the blueprint. The initial stage was conducted to gather comprehensive data about, in general, what is needed to be formulized in the blueprint. In this case, the data was gathered through interview and document analysis. Based on the data, the development stage could be started.

When referring to ADDIE (Analysis, Design, Development, Implementation, and Evaluation) model (see: e.g. Nita et al., 2022), the analysis and design stages have been covered in the previous process. This focuses on the development of the blueprint. However, it is important to note that in developing a blueprint, there are important steps to take into account before getting the best version of the document. As suggested by Fulcher (2010) it is important to conduct a prototyping and piloting. Prototype refers to small 'portion' of test items, developed based on the blueprint, to be tried out to small group(s) of prototyping subjects. When ready, it can continue to larger group(s) of individuals through piloting step. Those are for investigating the usefulness and usability of the blueprint (Fulcher, 2010). It means, when revisions or adjustments are needed, the blueprint should not be implemented in the real test development until the best version has been achieved. Nonetheless, the discussions on the prototyping and piloting are subjects to be conducted next.

4. Results and Discussion

In general, the main output of this study is a blueprint of EPT Polije to be implemented as the guideline for developing the English proficiency test. As explained previously, the focus of this study is developing the blueprint based on the analysis stage conducted prior to the development. The overview of the analysis stage is presented briefly followed by detailed elaboration regarding the development process.

4.1. Overview of the Analysis Stage

To initialize the whole process of the development of the blueprint, an analysis stage was conducted. An interview with the head of Language Center of Polije and document analysis of related documents including the ELT materials, specifically for the Intermediate English subject, the existing EPT blueprint, analysis of existing test package, and webpages informing recent development at Polije regarding the fields of study offered at the higher institution. The purpose was to gather substantial data about what aspects to develop in the new blueprint of EPT Polije. Additionally, focus group discussion (FGD) was also conducted with the head of Language Center of Polije, besides the informal discussions during the whole process, to design and develop the blueprint.

The results of the analyses showed that there are essential aspects to consider for the development of the blueprint of EPT Polije. Based on the interview, the general specifications of the blueprint could be formulated. This covers the format of the test, the skills to test, the parts in each section, and the number of items. Besides, through the interview with the head of Language Center of Polije, the details were also discussed. This was related to the CEFR levels to be distributed in the test and the scoring system.

The data gathered form the interview was, then, collaborated with the data from document analysis. This focused on reviewing the existing blueprint, analysis of the existing test package, ELT materials at Polije, and the vocational topics to be included in the test. In addition, various adjustments were also made through informal discussions. For instance, it was originally decided to develop of blueprint of the number of 200 test items. However, after careful consideration, especially regarding test taker characteristics, the total number was revised into 150 items by reducing 50 items from the listening section. Additionally, it was also decided that the English proficiency test only focuses to test at A2 to C1 levels of CEFR. Level A1 is not included because the target test takers are college students who have finished studying English at primary and secondary education. There are small amount of C2-level items to be tested to accommodate test takers with higher level of proficiency. Finally, the analysis as conducted in the previous stage was used as the basis of the development of the blueprint as presented as follows.

4.2. Development of Blueprint of EPT Polije

What it means by development of the blueprint of EPT Polije is the process of constructing the blueprint document. The blueprint is developed based on the data gathered from the previous stage, analysis. The following is the elaboration of the development.

4.2.1. General Information of the Blueprint of EPT Polije

The general information of the blueprint of EPT Polije was formulized to provide the big picture of how to complete the development of the blueprint and to what direction the blueprint should be constructed. In the document, this provides the general information about the EPT Polije as the test administered by the Language Center of Polije. The pieces of information included regarding the general information covers the identity of the document, the format of the test, test sections, parts of the test, number of items of each part, and total number of the test items.

In more details, as seen in the document, the official name of the test is English Proficiency Test (EPT) Politeknik Negeri Jember which is administered by Politeknik Negeri Jember, specifically Unit Penunjang Akademik (UPA) Bahasa which is also called Language Center of Polije. The format of the test is multiple-choice. Regarding the sections of the test, there are two including Listening and Reading. The Listening section is divided into three four parts: 1) Pictures Identification that covers 10 test items, 2) Questions – Responses that covers 15 test items, 3) Conversations which covers 15 test items, and 4) Talks that covers 10 items. Totally, the Listening Section has 50 items to be tested. Next, the Reading Section consists of two parts: 1) Vocabulary and Grammar which covers 40 test items, and 2) Reading Comprehension which covers 60 test items. It means that the total number of items to be tested in the Reading Section is 100. Overall, the Listening and Reading sections combined, there are 150 test items included in the blueprint. The following table 1 shows the general information as elaborated above.

Table 1. General Information of the Blueprint of EPT Polije

Blueprint of English Proficiency Test (EPT) Politeknik Negeri Jember					
Test Name	**Format**	**General Specifications**			
English Proficiency Test (EPT) Politeknik Negeri Jember	Multiple Choice	**Test Sections**	**Parts**	**Number of Items**	**Total**
		Listening	Pictures Identification	10	50
			Questions – Responses	15	
			Conversations	15	
			Talks	10	
		Reading	Vocabulary and Grammar	40	100
			Reading Comprehension	60	
		Total number of test items			150

4.2.2. Listening Section

The first section of the blueprint being developed was the Listening Section. As explained previously, there are four parts of this section including Pictures Identification, Questions – Responses, Conversations, and Talks. The brief information about this section can be seen in the Table. 1.

As briefly explained, the initial plan regarding the Listening Section was covering the total number of 100 items. However, through the FGD with the head of Language Center of Polije, it was decided to reduce the number into 50. The reason behind the decision was based on two aspects of consideration: 1) time allotment for the section and 2) concerns about students' readiness to face such a long duration the section. Those were based on professional justification, mainly, from the head of the Language Center who has administered the test, prior to the development of the new blueprint, in recent years. She explained that the recent version of the test only covered 24 listening test items. Therefore, having too many items could be hard both for the test takers and the Language Center who administers the test.

From the total of 50 test items, the section is divided into four parts. The first is pictures identification with 10 test items. The part is intended to test takers' ability to identify pictures and choose a spoken statements that best describe the pictures from multiple options. It means that what the test takers will hear are 10 sets of statements from which they should choose one for each item. In this case, the pictures will be divided into 7 categories: 1) two items of identifying object in a picture, 2) two items of identifying people's appearance in a picture, 3) two items of identifying an activity in a picture, 4) one item of predicting context in a picture, 5) one item of predicting relationship between people in a picture, 6) making mild inference about a picture, and 7) interpreting a scene or workplace task in a picture. From those, regarding the CEFR levels, category 1, 2, and 3 are at A2, category 4, 5, and 6 are at B1, and category 7 is at B2. It is considered that this part does not require complex cognitive and analytical ability, therefore no C1 nor C2 levels of CEFR tested.

In the next part, the test takers will be dealing with Questions – Responses part. This part consists of 15 test items in which the test takers should identify a response of a question for each item. In this case, the communicative ability being tested is how to response a question correctly in the given context. For instance, if the test takers hear, "Would you mind showing me the new flyer?" the expected response to the question is "Not at all." In this case, the item is intended to test the test taker's awareness that the question should be response in negative way or expression to show the agreement. Regarding the CEFR levels, 3 items are set

at A2, 4 items are at B1, 5 items are at B2, 2 items are at C1, and 1 item is at C2. In addition, for the first two parts, the levels of CEFR are distributed gradually. It means that the lower levels are set prior to the higher ones.

The part following the previous is conversations. Totally, there are 15 test items that belong to the part. This consists of 5 conversations with 3 questions for each. In this case, the focus of the part is to test the ability to comprehend longer spoken texts in forms of conversations. The test takers are expected to demonstrate their ability in understanding different aspects including the detailed information, either explicit or implicit, the topic of the conversations, and other aspects as suggested in the CEFR. It is important to note that there are several anchor items in this part. Those are items of the questions asked about the same matter. In this case, the anchor question is, "What is the conversation about?" and "What are the speakers discussing?" which have the same intention. The anchor questions are used to provide the consistency needed when developing the test especially when it comes to developing multiple sets of test packages. About the CEFR levels, each conversation consists of different levels of difficulty. Yet, those are distributed gradually based on the conversations presented. As an illustration, Conversation 1 consists of 1 item at B1 and 2 items at B2, while Conversation 2 consists of B1, B2, and C1 questions. This goes to the last conversation which consists of B2, C1, and C2 – level questions. Out of 15 questions, 3 items belong to B1, 7 items belong to B2, 4 items belong to C1, and 1 item belongs to C2.

The last part of the Listening section is Talks. The part is intended to test the test takers their ability in understanding longer talks in English. There are 3 talks presented in this part. Talks 1 and 2 have 3 test items each while Talks 3 consists of 4 questions. The distribution of the CEFR levels is similar to the previous section. Each talks accommodate different levels of difficulties ranging from B1 to C2. Out of 10 items, this section has 2 items at B1, 5 items of B2, 2 items at C1, and 1 item at C2.

Besides the details explained above, it is also important to inform that the blueprint also consists of the suggested type and/or duration of the audio or recordings. For instance, audio for the first part, Pictures Identification, is presented as the options to choose. For the second part, Questions – Responses, the recordings are presented for both the questions and responses, from which the test takers should choose one option. For the last two parts, the test takers will hear the recordings of the conversations and talks as well as the questions. These also indicate how long the recordings will last. Besides, specifically for the Talks part, the durations of the talks are indicated in the blueprint explicitly ranging from 60 – 90 seconds for Talk 1 to 120 – 160 seconds for Talk 3.

4.2.3. Reading Section

The Reading section of the EPT Polije designed in the new blueprint comprises 100 test items. This section has two parts: 1) Vocabulary and Grammar and 2) Reading Comprehension. The first part covers 40 test items and the second one covers 60 test items. The details are presented as follows.

The first part of the Reading section is Vocabulary and Grammar. As suggested by the term used, this is to test the ability related to vocabulary mastery and the grammar of English. This part covers 40 test items in forms of sentence completion. In this case, the test takers will need to identify incomplete sentences and should choose the best options available to complete them. Out of 40 test items, 8 are set at A2 level of CEFR, 15 items at B1, 11 items belong to B2, 4 items at C1, and 2 items at C2. The word-count for each level is also formulized. For the A2 level, each sentence ranges from 10 to 15 words, while the B1 should be between 15 and 25 words. This goes gradually to around 45 words for the C2 level. Additionally, the distribution of the CEFR levels also indicates the complexity of the sentences. It starts from

simple sentence with high- frequency words for the A2 to dense compound-complex sentences with layered clauses consisting of highly precise vocabulary for the C2 level. To demonstrate, for the A2 level, the test taker will read:

"Bali is one of the most _____ destinations in Indonesia."

While for the B2, they will read:

"Rapid development in information technology has influenced many sectors including agriculture. _____, the integration of internet of things (IoT) technology in this sector is crucial."

As expected, the first sample is a simple sentence that is intended to test the ability regarding parts of speech in English. For the B2 level, the test takers are expected to identify and choose the most appropriate discourse marker.

The next part is Reading Comprehension. In this part the test takers will have to read functional texts and answer the questions related to them. The questions cover the ability to identify general understanding towards a text such as about the topic, main idea, and purpose of a text, identify explicit information in the texts, identify implicit information of the text, and identify more complicated aspects of the text based on their levels of difficulty. There are 19 sets of functional texts followed by 2 to 6 questions each depending on the intended difficulties. What it means by 'set' is the fact that the test takers will have to read not only single texts to answer related questions but also paired texts to do so. As in illustration, the first passage is a short functional text such as advertisement followed by 2 questions. The passage 16, which is intended to be more difficult based on the CEFR levels, consists of a pair of email and press release, for example, from which the test takers should answer 4 questions.

As mentioned previously, this part comprises 60 test items. In this case, the levels of CEFR are distributed gradually, similar to the ones in Conversations and Talks parts in the Listening section. For instance, the first passage covers A2 and B1 levels of CEFR while the last passage, which is the most difficult, covers B2 to C2 levels of CEFR. Out of 60 items, there are 4 that belong to A2 level, 10 items at B1, 19 items at B2, 25 items at C1, and 2 items for the C2 level of CEFR. As implied, this part is intended to be more difficult from the previous one as indicated by the distribution of the CEFR levels; more C1-level questions are given in this part. For the A2 level, for instance, the test takers will answer a question about the purpose of a text. For the B2 – C1 levels, the test takers may have to deal with a question about making inference based on paired texts which require higher analytical skill.

4.2.4. CEFR in the Blueprint of EPT Polije

One of the key aspects of developing the blueprint of EPT Polije is incorporating the CEFR into the blueprint as the core reference. This required careful process starting from formulating the distribution, regarding the percentage of each level, to deciding what to test for each test item according to the CEFR. In this case, the main reference used in this study is *British Council - EAQUALS Core Inventory for General English* (North et al., 2015). However, frankly speaking, it is premature to claim that the blueprint resulted from this study is perfect. Further review and test should be made to confirm the validity and reliability of the blueprint especially regarding the implementation of the CEFR.

The discussion, regarding this matter, started by deciding the portions of CEFR levels in the blueprint. By considering various aspects, mainly based on the data gathered in the

analysis stage, it was decided that the blueprint focuses on A2 to C1 levels. The A1 level was eliminated because the test will be administered at university level where students mostly have passed their secondary education in which they learn English. However, the C2 level is still presented in a few number of test items to accommodate test takers with higher proficiency.

In the Listening section, out of 50 test items, 9 items are at A2 level or 18% from the total number, 12 items or 24% are at B1, 18 items or 36% are at B2, 8 items or 16% are at C1, and 3 items or 6% are at C2 level of CEFR. One the other hand, the Reading section consists of bigger number of test items. Out of 100 items, 12 items or 12% are at A2, 25 items or 25% are at B1, 30 items or 30% are at B2, 29 items or 29% are at C1, and 4 items or 4% are at C2 levels of CEFR. Combined altogether, the A2 level covers 21 items or 14% of the total number, B1 covers 37 items or 24,6% of the total number, B2 covers 48 items or 32% of the total number, C1 covers 37 items or 24,6% of the total number, and C2 level covers 7 items or 4,6% of the total number. The figure 1 bellow shows the brief picture the distribution of the CEFR levels.
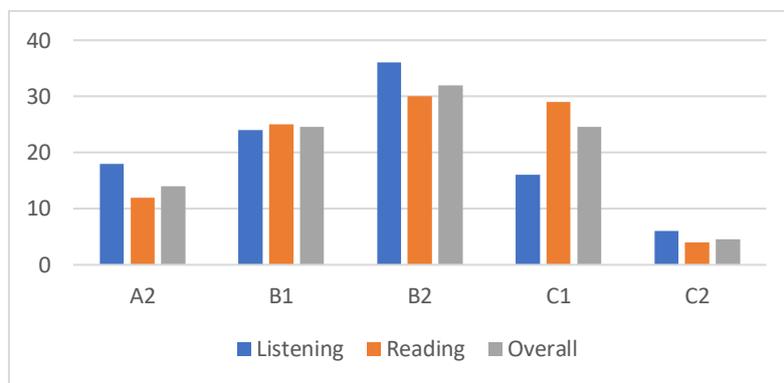

Figure 1. Distribution of CEFR Levels (in %)

As mentioned above, the main reference regarding the CEFR is British Council - EAQUALS Core Inventory for General English (North et al., 2015). This provides the guidelines on what English skills and components covered by each level along with various examples. Besides, the CEFR Descriptors by Council of Europe (available at: https://www.coe.int/en/web/common-european-framework-reference-languages/cefr-descriptors) was also used as supporting reference especially related to the skills to be tested.

4.2.5. Vocational Nuance in the Blueprint of EPT Polije

Another important aspect to consider in developing the blueprint of EPT Polije is the vocational nuance. This is related to the disciplines or fields of study to be included in the test. Mainly based on the results of analysis stage, the disciplines fall into 6 categories: 1) Agriculture, 2) Information Technology, 3) Engineering, 4) Health, 5) Language, Communication, and Tourism, and 6) Business.

In the implementation, those are grouped into 3 clusters based on how common topics under the fields of study are discussed in English especially at Politeknik Negeri Jember. Based on the discussion and by referring to the reference mentioned previously, the clusters can be seen in the table 2.

Table 2. The Clusters of Fields of Study

Cluster	Fields of Study
A2 – B1	Language, Communication, and Tourism
B1 – B2	Engineering; IT; Health
B2 – C1	Agriculture; Health; Business

The purpose of clustering the fields of study, as seen in the table 2 above, is to create efficient, systematic, and uniformed method when it comes to constructing multiple packages or sets of test. As an illustration, when the team of test developers need to create or find functional text for Reading Comprehension part at B1 level in form of advertisement, the text should be related to Engineering, IT, or Health. For sure, this will need mutual perception and agreement in practice while the other aspects like the word count and complexity should also be considered. However, it is important to remember that the clusters are assumption-based. Therefore, further studies might be needed to achieve more scientific justification.

4.2.6. Scoring and Conversion

Based on the interview with the head of Language Center of Polije in the analysis stage, a crucial adjustment to make in the new blueprint of EPT Polije is the scoring system. In this case, it is expected that the EPT scores achieved by the test takers are easy to interpret, especially when they need to compare the results with the scoring system of globally recognized English proficiency test.

By considering such need, a new scoring system was formulated. For the Listening section, each raw score is valued 10 points for the conversion. The maximum score for this section is 500. It means that when a test taker can answer 35 questions correctly, he/she gets 350 score. However, the conversion for the Reading section is quite different. This is due to the fact that the section has 100 test items. In this section, each raw score is valued 5 to get the maximum of 500 score when all questions are answered correctly. It means that when the same test taker can answer 70 questions correctly, he/she gets 350. The EPT Polije score is achieved by combining scores from both sections. As a result, when it comes to the previous illustration, the test taker gets the total of 700 as his/her EPT score.

The next, tricky, step was how to interpret the scores based on the CEFR. Again, through careful discussion, it was decided that the comparison is based on the percentage of each level distributed in the test. This is based on the assumption that if a test taker's real proficiency is at around B1, he/she might be able to complete the B1-level questions only. So do the other individuals with different levels of proficiency. To illustrate from the previous, the test taker who gets 700 has the ability between B1 and B2. Refer to Table 3 for the score mapping.

Table 3. Mapping Table of Scores and CEFR Levels

CEFR Level	Listening Scores	Reading Scores	Overall Scores
C1	470	480	950
B2	400	335	735
B1	210	185	395
A2	90	60	150

4.3. Discussion

The previous part of this paper has provided the insights into the process of developing the blueprint of EPT Polije. To get to this point, two stages have been done including the analysis stage and the development of the blueprint of EPT Polije. The long process is in line with Fulcher's (2010) explanation which says that developing a standardized test can take long time and very expensive. The process of the development usually requires a team who simultaneously work for the project as similar to what Cheewasukthaworn (2022). Bachman and Palmer (2010) even state that selecting a team involved in the process is one of major steps. It is confirmed that careful process should be undergone for expected quality of the test. As added by Fulcher (2010), "… how important tests are for the lives of the test takers …."

The purpose of the development of the blueprint is to accommodate the needs of English proficiency test in vocational context. The test takers coming from this area may have specific characteristics and needs compared to those in the academic ones. This is in line with the argument by Ojochegbe (2024) who explains that a test should take into account the cultural and linguistic diversity of the test takers. In this case, China seems to have stepped ahead. For about two decades, the have administered Practical English Test for Colleges (PRETCO), a standardized English test designed specifically for test takers coming from vocational and technical colleges (Chen, 2022), which is specifically designed to focus on the test takers' English proficiency in everyday and career related communication. This should be considered as the model demonstrating the importance of developing English proficiency test for specific needs and contexts. This is one of the big reasons why the project is conducted.

In developing the blueprint, CEFR is incorporated. Surely, this matter is nothing new in the development and administration of English assessments. As explained by Piccardo (2019), CEFR provides the basic foundation in English language assessment, including standardized tests, to get the full picture of test takers' proficiency. An important issue to consider regarding the application of CEFR in developing standardized test is that the interpretation may vary between institutions or stakeholders. The different formats and scoring systems are considered to cause the different interpretation (Waluyo et al., 2024). Therefore, it is important to state that the blueprint of EPT Polije may not provide similar conclusions regarding the scores and claim when compared to other similar tests.

5. Conclusion

The study has resulted the blueprint of EPT Polije. The development process does not stop at this point. The development will continue to the next step: prototyping and piloting. By conducting those to measure the quality of the design especially regarding the validity and reliability. When revisions needed, it is important to make necessary adjustments before performing the real test.

5.1. Acknowledgments

6. References

Arnelli, A. (2025). Students ' Perceptions on the TOEFL Test and Its Relevant Factors. *Educare: Journal Educational and Multimedia*, *03*(02), 221–228.

Azhari, T., Sahputri, J., & Rasyimah, R. (2021). Analysis of Student Perceptions on the Implementation of Required TOEFL Test before Graduation. *Proceedings of the 2nd International Conference on Social Science, Political Science, and Humanities (ICoSPOLHUM 2021)*, *648*(ICoSPOLHUM 2021), 125–130.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.

Cheewasukthaworn, K. (2022). Developing a Standardized English Proficiency Test in Alignment with the CEFR. *PASAA*, *63*(1), 66–92.

https://doi.org/10.58837/CHULA.PASAA.63.1.3

Chen, Y. (2022). PRETCO : An English Test for Vocational and Technical College Students. *International Journal of TESOL Studies*, *4*(1), 213–227.

Council of Europe. (2025). *Common European Framework of Reference*. https://www.coe.int/en/web/common-european-framework-reference-languages/introduction-and-context

Davidson, F. (2012). *Test specification and criterion referenced assessment* (G. Fulcher & F. Davidson (eds.)). Routledge.

Fulcher, G. (2010). *Practical Language Testing*. Hodder Education.

Nita, F. R., Astiandani, F. R., Wicaksono, A. L., & Janah, K. E. N. (2022). Using ADDIE model to develop learning materials of the test of English proficiency in Edmodo Using ADDIE model to develop learning materials of the test of English proficiency in Edmodo. *EnJourMe (English Journal of Merdeka): Culture, Language, and Teaching of English*, *7*(1), 62–77. https://doi.org/10.26905/enjourme.v7i1.7036

North, B., Ortega, A., & Sheehan, S. (2015). *British Council – EAQUALS Core Inventory for General English A Core Inventory for General English* (2nd ed.). British Council/EAQUALS. https://www.eaquals.org/wp-content/uploads/EAQUALS_British_Council_Core_Curriculum_April2011.pdf

Ockey, G. J., & Gokturk, N. (2019). Standardized Language Proficiency Tests in Higher Education. In *Springer International Handbooks of Education: Vol. Part F1628* (Issue January, pp. 377–393). https://doi.org/10.1007/978-3-030-02899-2_25

Ojochegbe, A. T. (2024). Rethinking Standardized Testing in English Language Proficiency : Moving Toward Culturally Responsive Assessment Models. *Jurnal Pendidikan Indonesia*, *5*(12), 1990–1996.

Piccardo, E. (2019). *TIRF language education in review –The Common European Framework of Reference (CEFR) in Language Education : Past , Present , and Future*. TIRF & Laureate International Universities. https://www.tirfonline.org/wp-content/uploads/2021/10/LEiR_CEFR.pdf

Roza, D. (2019). The Challenges and Strategies in Teaching TOEFL and IELTS Test Preparation. *J-SHMIC : Journal of English for Academic*, *6*(2), 1–13.

Susyla, D., Angraini, R., & Ananda, R. P. (2021). The Lecturers ' English Ability Through The TOEFL Test. *Linguistic, English Education and Art (LEEA) Journal*, *5*(1), 69–82.

Waluyo, B., Zahabi, A., & Ruangsang, L. (2024). Language Assessment at a Thai University : A CEFR-Based Test of English Proficiency Development. *REFLections*, *31*(1), 25–47.

Yasmin, F. (2024). Redefining Language Proficiency Testing : Addressing Cultural Biases and Proposing Comprehensive Solutions . *International Journal of Novel Research and Development (IJNRD)*, *9*(3), 408–422.