

A Review of the Test Items of the LPATE From the Perspective of the CEFR's 'Can-Do' Descriptors

Tantri Sari Safitry
University of Indonesia and Politeknik Negeri Jakarta

Untung Yuwono
University of Indonesia

Sisilia Setiawati Halimi
University of Indonesia

This research was conducted to analyze the content validity of the Language Proficiency Assessment for Teachers (English Language) (LPATE) adapted to the 'can-do' descriptors of Common European Framework Reference for Languages (CEFR). Through descriptive qualitative research, the assessment questions were analyzed using the validity formula of Suharsimi Arikunto and Pearson. The results showed that the LPATE has good validity and is following the CEFR's 'can-do' descriptor. Although some of the items discussed things that were not following the 'can do' statement on the CEFR, these did not interfere with the validity of the test because the material and types of questions given were in line with the final competency achievement of the LPATE.

Keywords: English proficiency assessment, LPATE test specification, contents validity, CEFR

INTRODUCTION

To measure learning outcomes accurately, teachers must use a good test. However, making a test called 'a good test' is not easy because test makers are asked to meet the existing requirements and characteristics (Aswal, 2020). In foreign language testing, the validity of the test is one thing that deserves careful attention, along with reliability, authenticity, practicality, interaction, and effect (Bachman & Palmer, 1996). Therefore, how to design a valid test is one of the concerns of researchers worldwide, including in Indonesia (Furwana, 2019; Aswal, 2020; Bachman & Palmer, 1996; Thao, 2018).

A study explains that the validity of a test is carried out to measure what should be measured and nothing else (Heaton, 1998). This study also demonstrates that without validity, there can be no guarantee that the conclusions obtained from the test results can be trusted. Wiersma & G. Jurs (1990) also added that validity is the extent to which the test measures what it wants to measure, whether it is achievement, talent, or proficiency in the language. If a test has high validity, the results will meet the criteria of a good test: reliability, authenticity, practicality, interaction, and effect (Bachman & Palmer, 1996). One of the validities

that can be studied is content validity. Heaton (1998) explained that content validity relates to previously learned material. To fulfill this, test makers usually must refer to the teaching syllabus.

In this study, the researchers have not yet referred to the teaching syllabuses used in existing English education study programs. Still, it has first referred to the Common European Framework of Reference of Languages (CEFR). This research is one of the initial studies of research on English language proficiency assessment for teachers, which will be examined concerning the syllabus applied in Indonesia. The researchers wanted to know if the contents of the English proficiency test for teachers is valid and per the standards set by CEFR. This is in line with the opinion that content validity can ensure that the measurement includes an adequate set of items and represents the concept expressed (Hendryadi, 2017). A test is also said to have content validity if its content has a representative sample of the language skills, structure, etc. are contained in the objectives that want to be achieved.

Many experts believe that if teachers want to assess the ability of foreign language learners objectively and comprehensively, especially in English, they must use a firm standard. Many international standards can be used to measure language level proficiency, such as The American Council on the Teaching of Foreign Languages (ACTFL) proficiency level (See for more: <https://www.actfl.org/>) or the Cambridge proficiency test (See for more: <https://www.cambridgeenglish.org/>). One of the proficiency levels is the Common European Framework Reference for Languages (CEFR hereinafter). The CEFR Has been used as a guideline to describe students' foreign language abilities (English is one of them) throughout Europe. Recently the framework has been adopted and modified as a benchmark for assessing the language proficiency of teachers and students in countries such as Malaysia, Thailand, and Vietnam (Foley, 2019; Nguyen & Hamid, 2020; Uri & Abdul Aziz, 2018). Lately, in Indonesia, there is also great interest from the experts to implement the framework into their works, like the use of CEFR in BIPA materials (*Bahasa Indonesia bagi Penutur Asing – Indonesian Language for Foreign Speakers*) (Sudaryanto & Widodo, 2020) or for teaching French in several universities (Renandya, Hamid, & Joko, 2018).

Moreover, along with the implementation of the Merdeka Belajar Kampus Merdeka curriculum (MBKM), based on Article 18 (b) Permendikbud Number 3 of 2020 concerning National Higher Education Standards, it is explained that the MBKM curriculum encourages students to master various competences that are useful for entering the world of work. Students can choose courses according to their needs. Then, one of the competencies that they must have to compete in the world of work is to have language skills, especially English. To compete in a global society, students are expected to have a certificate of proficiency in English, especially for prospective English teachers. Quoted from Global.exam.com, the scores of several general English proficiency tests tested in Indonesia, such as the TOEFL (Test of English as a Foreign Language) and IELTS (International English Language Testing System) have been converted to proficiency levels in the CEFR framework. Thus, it can be concluded that a good proficiency test is expected to be associated with the CEFR framework.

The use of international proficiency standards in determining student's English proficiency is significant so that language competencies can be truly measured. Therefore, the researchers wanted to find out whether the contents of the LPATE were following the standard description from the CEFR. Suppose the results of the study show the suitability of the content validity of the LPATE with the CEFR description. In that case, the results of this study can be used as a justification to see the suitability of this language proficiency test for English teachers with the syllabuses used in English Education Studies Programs in Indonesia. In Indonesia, quoted some Heads of the Department of English Education Language Study Program, the minimum score for the PBT TOEFL test that the students must have if they want to graduate from the campus ranges from 450 to 500 or 100 for iBT TOEFL test. This statement aligns with Renandya, Hamid, & Joko (2018), who stated that most universities in Indonesia set the score range from 450 to 525 for PBT TOEFL test. If we convert the score to CEFR, their scores will be on the B2 CEFR level (See <https://www.paradigm-edu.com/toefl-pbt/> for more).

The language proficiency assessment for English teachers used in this study was taken from Hong Kong by Language Proficiency Assessment for Teachers (English Language) (LPATE). The LPATE has been used since 2001 by the Hong Kong Examinations and Assessment Authority (HKEAA) and the Education Bureau (EDB) to evaluate English language skills for English teachers at primary and secondary levels, as

well as graduates of English programs who wish to teach in Hong Kong (Coniam & Falvey, 2018). The LPATE is divided into several language skills: reading, listening, writing, speaking, and a Classroom Language Assessment (CLA hereafter). In the CLA, the participants are assessed in terms of their (a) grammar and linguistic accuracy and range, (b) pronunciation, stress, and intonation, (c) the language of classroom interaction, and (d) the language of classroom instruction. Those components are later scored on a five-level scale with the mid-point (level 3) as the benchmark. On their paper, Renandya, Hamid, & Joko (2018) stated that the Director of the Testing Center LPATE, Mr. Drave, N., that level 3 on the LPATE is equivalent to IELTS 7 or CEFR C1.

Since there is no particular language proficiency test for English teachers (Renandya, Hamid, & Joko, 2018), the LPATE is considered capable of being one of the standardized proficiency tests that can be used in Indonesia. However, because it has not been found –to avoid mentioning that there are no previous studies–, thus, a study that tries to analyze the validity of the content of the language proficiency test from the CEFR’s point of view is considered necessary. Based on this background, the research question in this study is “Is the Language Proficiency Assessment for Teachers (English Language) from Hong Kong (LPATE) in line with the CEFR standard?”.

A plethora of research that focuses on language proficiency tests has been carried out in several studies (Thao, 2018; Yan & Huizhong, 2006; Huang & Flores, 2018; Al Malki, 2014; Tran, Griffin & Nguyen, 2010; Rachmawati & Purwanti, 2021; Nugroho, 2018; Aniroh, 2009). A critical evaluation was carried out on the language proficiency test for the 21st century, which was used to test English proficiency for eight member states in the United States; from the observations, it was found that there is still a lack of research results validity of the test so that it raises doubts on the quality of the test given (Huang & Flores, 2018). In China, the College English test (CET) was carried out as well as an in-depth review of the content of the test. The result showed that the CET test still had to develop the content of the test so that it was following communicative competences that are needed to meet the ever-changing needs of society (Yan & Huizhong, 2006). Besides some research on the proficiency test itself, a case study was also conducted by Rachmawati & Purwanti (2021) to see the effect of English teachers’ proficiency on the teachers’ classroom instruction capability. The result shows that all teachers must obtain both high proficiency levels and pedagogical skills to support the teaching learning activities. Kostina (2012) in her research, also revealed that only a few pre-service teachers only achieved B2 level. Therefore, it is strongly recommended that pre-service teachers strengthen not only classroom communication and evaluation practices but also implement curricular changes. Nugroho (2018) also researched the proficiency of English teachers at the junior and senior high school levels who joined Pendidikan dan Pelatihan Profesi Guru (PLPG) at the State University of Surabaya in 2017. The results showed that most participants stated they were confident with their competencies but also needed continuous self-development activities, specifically for their language skills. Various activities are proposed. Some of their proposals are to maintain the four micro-skills, provide scholarships for teachers to take short course programs abroad, and maximize the function of the subject teachers association or *Musyawarah Guru Mata Pelajaran* (MGMP). Meanwhile, Aniroh (2009) discussed the need for teachers in the field of English for Specific Purposes (ESP) to have good quality teaching, one of which is proficiency in English. However, she did not discuss the issue of skills in more detail.

Even though the role of CEFR in the development of language teaching and learning does not escape debate among researchers (Alderson, 2017; Brunfaut & Harding, 2020), they continue to adopt CEFR not only on the syllabus, curriculum guidelines and language textbooks but also on their language proficiency test (Council of Europe, 2001). The CEFR comprehensively describes what language learners must learn to use language to communicate. CEFR also can help learners acquire the knowledge and skills that must be developed to use the language effectively. In addition, this framework can also define a language learner’s level of proficiency that can be measured at each stage of learning. The framework also supports research and development initiatives in the education sector that introduce the most suitable techniques and resources to assist different classes and types of students in acquiring communicative skills that suit their individual needs at all levels of education. In addition, the CEFR describes how good a person is at reading, writing, listening, speaking, and understanding a foreign language (Quoted from efset.org, 2022). With six foreign language proficiency levels: A1 (Breakthrough), A2 (Waystage), B1 (Threshold), B2 (Vantage),

C1 (Effective Operational Proficiency), and C2 (Mastery), CEFR provides a set of ‘can do’ descriptors that list functions that a person can do at any given level.

For example, one can-do descriptor at level A1, the breakthrough level, is ‘can understand and use common everyday expressions and fundamental phrases aimed at meeting the need for concrete types’. It is hoped that a foreign language teacher can use this statement to evaluate skills and then design appropriate lessons to address their knowledge gaps. Meanwhile, level A2 does show up to reflect the level alluded to by the way stage specification. On this level, most descriptors are expected to express social functions like simple everyday polite forms of greeting and address: greet people, ask how they are and react to the news, handle very short social exchanges, and more. B1 level, the threshold one, is categorized by two features. The first feature is the ability to maintain interaction and get across what you want in various contexts. The second one is the ability to cope flexibly with problems in everyday life. Next, the B2 level is intended to reflect the vantage level specification. Within this level, the learners are already slow but sure joining the upper intermediate level, which means metaphorically speaking that they have arrived abroad somewhere, and things look different; thus, they acquire new perspectives and can look around them in a new way. For example, at the lower end of the band, there is a focus on an effective argument: account for and sustain his/her opinions in discussion by providing relevant explanations, ideas, and comments; explain a viewpoint on a topical issue giving the advantages and disadvantages of various options; construct a chain of reasoned argument; develop a statement giving reasons in support of or against a particular point of view, and more.

Meanwhile, the following band, the C1 was named as effective operational proficiency. This level is good for providing fluent and spontaneous communication, and the last is C2 level. Despite being referred to as “Mastery,” C2 level does not necessarily indicate proficiency in a native or close to the native language. The goal is to describe the level of clarity, appropriateness, and comfort with language that characterizes the speech of people who have achieved great achievement as language learners (Council of Europe, 2001).

For reading skills, the competence has five parts for each ‘can-do’ descriptor. First is Reading a text “diagonally” while quickly deciding whether to read (parts of it) properly is known as Reading for Orientation (skimming). Second, Reading for Information or argument involves scanning a text fast in pursuit of a particular item, usually a piece of information (scanning). Next, Reading as a Leisure Activity means reading involves non-fiction, but not literature. It will also encompass magazines and newspapers, blogs, biographies, etc. – and possibly even texts another person would read only for work or study purposes, depending on one’s interests. Fourth, Reading Instructions is a specialized form of reading for information. Lastly, Reading Correspondence is when the scales start in each category with interpersonal language use.

The written production skill is divided into two categories, namely Creative Writing and Written Reports and Essays. For the former, it involves personal and imaginative expression in various text types, while the latter covers more types of transactional and evaluative writing. Meanwhile, the written interaction scales are Correspondences and Notes.

For reception activities, the scales are divided into several parts in listening skills. The first one understands the conversation between other speakers. It concerns two main situations: the first is when other speakers in a group interaction talk across the user/learner to each other so that the user/learner is no longer directly addressed. The second situation is when the user/learner is an overhearer: listening to a conversation between people nearby. The second one is Listening as a member of a live audience. It deals with listening to a speaker addressing an audience, for example, in a meeting or seminar, at a conference or lecture, on a guided tour, at a wedding or other celebration. Next is Listening to announcements and instructions; it involves a different type of highly focused listening in which the aim is to catch specific information. Another one is Listening to audio media and recordings; it involves broadcast media and recorded materials, including messages, weather forecasts, narrated stories, news bulletins, interviews, and documentaries.

Spoken production in CEFR ‘can-do’ descriptors is consisted of several scales. There are Sustained Monologues for Describing Experience, Giving Information, Putting a Case (like in a debate). Moreover, the ranking also has the Public Announcement and Addressing Audiences. On the other hand, the spoken

interaction scales are organized by three macro-functions: interpersonal, transactional, and evaluative. The descriptors are divided into Understanding an Interlocutor, Conversations, Informal Discussions with friends, Formal Discussions (meetings), Goal-Oriented Co-Operation (e.g., cooking together, discussing a document, organising an event etc.), Obtaining Goods and Services, Information Exchange, Interviewing and being interviewed and Using telecommunications.

Based on the importance of the CEFR, this research was aimed to analyze the content validity of the Language Proficiency Assessment for Teachers (English Language) (LPATE) adapted to the ‘can-do’ descriptors of Common European Framework Reference for Languages (CEFR). As stated before, this was served as one of the initial studies of research on English language proficiency assessment for teachers, which will be examined concerning the syllabus applied in Indonesia.

METHOD

The data in this study were obtained through documentation, listening, and notes. The data source of this research is the ‘can-do’ descriptors on the LPATE in the form of test items with a total of 91 questions (see Appendix 1). The researchers analyzed the test items and whether each item on the LPATE represents the indicators as stated in the CEFR, which is used as a reference to measure the validity of the contents of the LPATE test. By using qualitative description, this study tried to describe the suitability and discrepancy of the LPATE test with the CEFR ‘can-do’ descriptors. There were two instruments used in this study. First, to analyze content validity, the researchers used the formula from Arikunto (1992) adopted from Pearson (2018). Second, the researchers also asked for help from experts in validating what was obtained from the instrument and the results were analyzed using Lawshe’s Content Validity Ratio (CVR) (1975).

To analyze the data, the researchers used descriptive analysis by the number of each item and then identified each item related to the test indicators. After that, the researchers compared the details of the data from the test items with the ‘can-do’ descriptors from the CEFR and then explained the validity of the contents of the LPATE qualitatively. To tabulate the data, the researchers changed the description of the CEFR into codes, for example, A1, B1, C1, and so on (see Table 2). Then, the items on the test were changed to numbers, such as 1, 2, 3, and so on. Furthermore, to get a presentation of validity, the researcher used the formula from Arikunto:

$$P = \frac{F}{N} \times 100$$

P = Percentage
F = Frequency
N = The amount of sample

Then, the researchers categorized the percentage of test suitability found into several levels of criteria, including:

76-100%= Fine

56-75%= Enough

40-55%=Less

<40%=Bad

Moreover, for the Lawshe’s Content Validity Ratio (CVR), $CVR = (ne - N/2) / (N/2)$, where CVR is the content validity ratio, ne is the number of panelists who answered “important”, N is the total number of panelists. This formula measures agreement among raters about the suitability of content validity on a particular test. Experts adopted the criteria given to respond to statements for each item of the question.

FINDINGS

As mentioned before, the data for this research is the ‘can-do’ descriptors on the LPATE in the form of test items with a total of 91 questions, which include reading, writing, listening, speaking, and pedagogical

class language skills. The number of questions consists of reading with 46 questions, writing with three questions, listening with 35 questions, speaking with three questions, and pedagogical class language with four questions.

The following table describes the total frequency of conformity of the test items of the LPATE with the CEFR ‘can-do’ descriptors.

TABLE 1
DATA ANALYSIS

Skills	Item Code Question	Data Analysis
	A1	<p>The reading section on the test items were analyzed as follows:</p> <ol style="list-style-type: none"> 1. 1. Test items 1, 5, 22, 23, 28, 29, 31, 37, 45, 46 represent the Reading for Orientation part. The descriptor refers to the B1 level, which is ‘Can scan longer texts in order to locate desired information, and gather information from different parts of a text, or from different texts in order to fulfil a specific task’ and level C1 in Reading for Information and Argument ‘Can obtain information, ideas and opinions from highly specialized sources within his/her field.’. 2. Test items 2, 3, 4, 7, 8, 9, 10 represent the Reading for Information and Argument part. The descriptor refers to the B1 level, which is ‘Can understand most factual information that he/she is likely to come across on familiar subjects of interest, provided he/she has sufficient time for re-reading’. 3. Test items 6, 18, 19, 26, 30 represent the Reading for Information and Argument. The descriptor refers to the B2 level, which is ‘Can understand articles and reports concerned with contemporary problems in which the writers adopt particular stances or viewpoints’. 4. Test item 32 represents Reading for Information and Argument. The descriptor refers to the B1 level, which is ‘Can identify the main conclusions in clearly signaled argumentative texts’. 5. Test items 11, 12, 13, 14, 15, 16, 17, 20, 21, 24,25, 27, 33, 34, 35, 36, 38, 39, 40, 41, 42 represent the Reading for Information and Argument part. The descriptor refers to the B2 level, which is ‘Can obtain information, ideas and opinions from highly specialised sources within his/her field.’. 6. Test item 44 includes in Reading for Information and Argument part, it consists of the C1 level which is ‘Can obtain information, ideas and opinions from highly specialised sources within his/her field.’.

Skills	Item Code Question	Data Analysis
Writing	A2	Test item 47 represents Written Reports and Essays part. The descriptors refer to the B2 level , which are ‘Can write an essay or report which develops an argument, giving reasons in support of or against a particular point of view and explaining the advantages and disadvantages of various options’ and ‘Can write an essay or report that develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail’.
	A3	Test items 48 and 49 cannot be included in any ‘can- do’ writing descriptors of CEFR, because these items are related to identifying errors in the text.
Listening	A4	Test items 50 to 64 represent Listening to Audio Media and Recordings . The descriptor refers to the B1 level , which is ‘Can understand the information content of the majority of recorded or broadcast audio material on topics of personal interest delivered in clear standard speech.’
	A5	Test items 65 to 74 represent Listening to Audio Media and Recordings . The descriptor refers to the B2 level , which is ‘Can understand recordings in the standard form of the language likely to be encountered in social, professional or academic life and identify speaker viewpoints and attitudes as well as the information content.’
	A6	Test items 75 to 84 represent Listening as a Member of a Live Audience . The descriptor refers to the B1 level , which is ‘Can follow in outline straightforward short talks on familiar topics, provided these are delivered in clearly articulated standard speech’.
Speaking	A7	Test item 85 represents the reading aloud section and this doesn’t fit into any ‘can-do’ descriptors.
	A8	Test item 86 represents Spoken Production: Sustained Monologue: Putting A Case (In a Debate) . The descriptors refer the B2 level , which are ‘Can develop an argument systematically with appropriate highlighting of significant points, and relevant supporting detail’; ‘Can develop a clear argument, expanding and supporting his/her points of view at some length with subsidiary points and relevant examples’; ‘Can construct a chain of reasoned argument’; and ‘Can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options’ and this test item is also included in Oral Production in the C1 level , ‘Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.’.

Skills	Item Code Question	Data Analysis
	A9	Test item 87 represents Spoken Interaction: Discussion (Meetings) . The descriptor refers to the B1 level , which is ‘Can follow in outline straightforward short talks on familiar topics, provided these are delivered in clearly articulated standard speech’ and this test item is also included in Oral Production in the C1 level , ‘Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.’.
Classroom Language Assessment	A10	Test item 88 represents Spoken Interaction: Understanding an Interlocutor . The descriptors refer to the A2 level , which are ‘Can understand what is said clearly, slowly and directly to him/her in simple everyday conversation; can be made to understand, if the speaker can take the trouble’; ‘Can generally understand clear, standard speech on familiar matters directed at him/her, provided he/she can ask for repetition or reformulation from time to time.’; ‘Can understand enough to manage simple, routine exchanges without undue effort.’ and the B1 level, which is ‘Can follow clearly articulated speech directed at him/her in everyday conversation, though will sometimes have to ask for repetition of particular words and phrases’; and the B2 level , which is ‘Can understand in detail what is said to him/her in the standard spoken language even in a noisy environment’.
	A11	Test item 89 represents the importance of accurate pronunciation, stress and intonation in speaking and this does not suit with any descriptor in the CEFR. However, all descriptors for speaking production and interaction that require accurate pronunciation, stress and intonation fit best with this test item.
	A12	Test item 90 represents Informal Discussion . The descriptors refer to the A2 level , which are ‘Can generally identify the topic of discussion around him/her which is conducted slowly and clearly.’; ‘Can exchange opinions and compare things and people using simple language.’; ‘Can make and respond to suggestions’; ‘Can agree and disagree with others.’; ‘Can give brief comments on the views of others.’ And Oral Production in the C1 level , ‘Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.’.

Skills	Item Code Question	Data Analysis
	A13	Test item 91 represents Spoken Interaction: Conversation . The descriptors refer to the B1 level , which are ‘Can start up a conversation and help it to keep going by asking people relatively spontaneous questions about a special experience or event, expressing reactions and opinion on familiar subjects’; ‘Can have relatively long conversations on subjects of common interest, provided that the interlocutor tries to support’ and Oral Production in the C1 level, ‘Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.’.

The researchers calculated the frequency of suitability and discrepancy of language proficiency test items for English teachers with ‘can-do’ descriptors on the CEFR (written with code A1-A13) based on the data from the analysis of the things above. Based on the data from the analysis of the things, the calculation of the suitability and discrepancy of the questions studied can be seen in Table 4.2 below:

TABLE 2
CONFORMITY AND INCOMPATIBILITY OF TEST ITEMS CALCULATION

NO	TEST CODE	TEST QUESTION NUMBER												
		1-46	47	48-49	50-64	65-74	75-84	85	86	87	88	89	90	91
1	A1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	A2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	A3	0	0	2	0	0	0	0	0	0	0	0	0	0
4	A4	0	0	0	0	0	0	0	0	0	0	0	0	0
5	A5	0	0	0	0	0	0	0	0	0	0	0	0	0
6	A6	0	0	0	0	0	0	0	0	0	0	0	0	0
7	A7	0	0	0	0	0	0	1	0	0	0	0	0	0
8	A8	0	0	0	0	0	0	0	0	0	0	0	0	0
9	A9	0	0	0	0	0	0	0	0	0	0	0	0	0
10	A10	0	0	0	0	0	0	0	0	0	0	0	0	0
11	A11	0	0	0	0	0	0	0	0	0	0	0	0	0
12	A1	0	0	0	0	0	0	0	0	0	0	0	0	0
13	A13	0	0	0	0	0	0	0	0	0	0	0	0	0

TABLE 3
ANALYSIS OF CONFORMITY AND NON- CONFORMANCE OF TEST ITEMS OF
THE LPATE AND CEFR DESCRIPTORS

NO	Test Question	Number Item Question	Total frequency
1	Suitability test item with <i>CEFR</i> can-do description	1-46, 47, 48-49, 50-64, 65-74, 85, 86, 87, 88, 89, 90, 91	88 test items
2	Incompatibility item question with <i>CEFR</i> can-do description	48-49, 85	3 test items

Then, the content validity formula presented by Arikunto and Pearson was used to calculate the suitability of the data that had been analyzed. The content validity presentation obtained is described as follows (Hwang & Flores, 2018).

$$P = \frac{F}{N} \times 100$$

$$P = 88/91 \times 100\%$$

$$P = 96\%$$

Therefore, it can be concluded that the proficiency test items have a content validity level that corresponds to the 'can do' statement from the CEFR framework. Meanwhile, based on the data obtained, the result of the discrepancy between the skills question and the CEFR descriptors is

$$P = 3/91 \times 100\%$$

$$P = 0.032\%$$

As for the results of the calculation using the Content Validity Ratio (CVR) from Lawshe (1975), the results show that the panelists (in this case two English lecturers with teaching experience for more than 10 years) agreed that the LPATE questions in the CEFR framework point of view were essential. The panelists filled in the questionnaire with 3 options: important (i), appropriate but less important (a), and not important (n).

TABLE 4
RESULTS FROM THE CONTENT VALIDITY RATION (CVR) FROM
LAWSHE (1975)

NO	TEST CODE	PANELIST 1	PANELIST 2
1	A1	i	i
2	A2	i	i
3	A3	i	i
4	A4	i	i
5	A5	i	i
6	A6	i	i
7	A7	i	i
8	A8	i	i
9	A9	i	i
10	A10	i	i

NO	TEST CODE	PANELIST 1	PANELIST 2
11	A11	i	i
12	A12	i	i
13	A13	i	i

$$CVR = (ne - N/2) / (N/2)$$

$$CVR = ((0/2)/2) - 0 = 0$$

This formula ranges between +1 and -1. Positive values show that at least half of the panelists rate items as important/essential. The greater the CVR from 0 means that the more “important” and the higher the content validity is. The result of this study is 0. Values 0 and above indicate that more than half of the experts agree. It indicates that the items on this test could be considered as good questions. The panelists agreed that measuring a teacher candidate’s English proficiency is important by looking at the CEFR description guidelines.

DISCUSSION

Based on the data on the tables and the explanations in the previous section, the overall language proficiency test questions for English teachers range from levels A2 to C1 in the CEFR description. Here, it is proved that the LPATE test covers up to the C1 CEFR level, as stated by the Director of Testing Center LPATE, Mr. Drave, N earlier in 2018 (Renandya, Joko & Hamid, 2018). Though the Director did not explain clearly if all sections have the C1 level, it is found in this research that only reading, speaking, and CLA sections cover the C1 level, while the highest level for listening and writing is the B2 level. This information can be used as a suggestion for the stakeholder who managed the test to pay more attention to the upcoming test items with the C1 level questions should be added for all sections of the test.

In addition, it is also found that all sections of the test have test items that align with the ‘can-do’ descriptors of the B2 level of CEFR. Since most English Language Study Programs in Indonesia have targeted the B2 level (TOEFL score test ranges from 450 to 525) as their graduation requirement (Renandya, Joko & Hamid, 2018) therefore, it can be concluded that the LPATE test is followed by the needs of English teachers’ proficiency in Indonesia. However, there is a concern that this B2 level can be applied globally in Indonesia. According to the Political and Economic Risk Consultant (PERC) survey, Indonesia’s education quality is ranked 12 out of 12 countries (Purnawanti, 2016). Further, taken from the Central Bureau of Statistics in 2020, the education inequality in 34 provinces in Indonesia shows that the highest educational inequalities are Papua (0.418), West Nusa Tenggara (0.299), and West Kalimantan (0.296), while the lowest areas are DKI Jakarta (0.156), Riau Islands (0.188), and Moluccas (0.196). Education inequality in Indonesia primarily arises from the curriculum, educational policies, educational facilities, application of information and communication technology in education, affordable education costs, education management, and human resources (Pramana, et al., 2021). Therefore, to reduce inequality problems in education, several aspects of education need to be improved. One is the students’ English language proficiency, especially for the English Language Study Program students in rural areas.

The findings of this review also indicate that the content validity for the LPATE test is 96%, along with its 0’s CVR score, suggests that the test can be called a good and valuable test and follows the CEFR’s ‘can-do’ descriptors (Bachman & Palmer, 1996). As Frost (2022) explains, if the test has high content validity, it fully covers the subject for the intended audience. Meanwhile, lower scores imply that the test does not cover all necessary aspects of the issue.

In the reading section, test items 1 to 46 were under the descriptions at levels B1 to C1. The test items given in this section use short texts as the reading source. On this part, it can be said that the ‘can-do’ descriptors from the LPATE item test are dominated by the Reading for Information and Argument and Reading for Orientation descriptors. Both descriptors are also known to help the candidates skim and scan so the reader can get specific information from the text correctly (Council of Europe, 2001). Besides that,

almost all test items have conformity with the CEFR descriptors. From these findings, it can be inferred that the test items assessed the candidate's reading comprehension ability from a short text with a relevant topic. Having excellent reading comprehension is something that teacher candidates must have. How will they teach the students later if they do not possess it well? As Dwiningtyas, et al. (2020) stated that teachers must have a reading comprehension skill as well as be able to combine some teaching reading strategies to make students become proficient and successful readers.

Test item 47, "Write a short talk explaining why healthy friendships are important for successful academic achievement and how teachers can help students get out of 'toxic friendships'. Write about 400 words." is following level B2. On the contrary, for items 48 and 49, no matching descriptors are in line with the test items. The given test items ask the candidates to detect and correct errors/problems from a text as well as do some corrections of errors/issues. Nevertheless, the error analysis part was not matched with any CEFR level. The researchers assumed this test item was essential to be included in the proficiency test because it is similar to what Little & Byram (2004) said, a teacher must understand the types of errors and errors made by students. The teacher should also know when and how mistakes can be corrected.

For the listening skill, it was found that there was a match between test items 50 to 84 with descriptions at levels B1 to B2. They used voice recordings from podcasts, radio interviews, and graduation speeches. It can be concluded that the proficiency test has met the requirements in the CEFR descriptors. According to Pearson Education Asia Limited (2010), cited by Taylor and Geranpayeh (2011), these test items were applied to follow academic-level instruction and participate in undergraduate education programs, including coursework and student life. However, it is suggested that this section should contain the C1 level questions since it is considered appropriate for entry to post-graduate programs. Nonetheless, the researchers advised that the test makers must add more the C1 levels for the section in the future.

Additionally, the speaking test in test item 85 for reading aloud, 86 for recounting an experience/presenting an argument, and 87 for joining the discussion with peers were included in speaking production and interaction in the B2 to C1 CEFR descriptors level. Yet, seeing from the descriptors, the proficiency test did not contain a sustained monologue for describing an experience, giving information, making a public announcement, and addressing an audience. The researchers assumed that since the test wants to focus on academic speaking, it focuses on speaking production and interaction instead stated from the Vanderbilt.edu (2022), academic speaking is suitable for students whose English-speaking skills need improvement. The objective of an academic speaking test is to check all aspects of speaking ability, especially in educational contexts. Moreover, Correia (2016) also explained that since assessing speaking is at the center of teaching-learning interaction, it is a challenging, time-consuming, and complex process that English teachers require. Therefore, the LPATE test only contains the academic English speaking for the vocal production and interaction. Still, the rest of the descriptors can be assessed later in the classroom language assessment.

The last section is the CLA. This section has four main categories for scoring the candidates; they are (1) Grammatical and Lexical Accuracy and Range; (2) Pronunciation, Stress, and Intonation; (3) Language of Interaction and

(4) Language of Instruction. The first and second categories are assessed for the teachers' speaking competency in this proficiency test. The researchers inferred that both types are automatically integrated into all speaking section test items descriptors. During the speaking test, the grammar part is used to see to what extent those assessed can use a wide range of English grammar accurately while speaking. In contrast, the vocabulary is used to check that the vocabulary selections of those considered are appropriate and rich in variety. At the same time, the pronunciation part tries to see to what extent the candidates can pronounce English accurately and naturally (Kiruma et al., 2017). Therefore, all the speaking production and interaction descriptors are following the test items.

On the other hand, the third and fourth categories were suitable with the A2 to B2 level for the Language of Interaction and Instruction descriptors. These descriptors are proposed to assess the candidates' English-speaking production and interaction competencies during their teaching activities. As stated by Kimura et al. (2017), the language of instructions is used to check whether the given instructions are efficient and

straightforward, while the language of interactions is tried to assess whether the candidate's interaction with students is going smoothly and effectively.

Another study from Nakata (2010) also discovered that the CLA's categories improved candidates' awareness of classroom English use and showed a strong potential to enhance their teacher language awareness (TLA). This TLA refers to language teachers' general knowledge and beliefs regarding language and the language they teach (Andrew, 2017). The TLA also bridges pedagogical content knowledge and proficiency in foreign language (like English). By having this language awareness, hopefully, the prospective teachers can facilitate their students' learning (Andrew, 2017) and create a thriving learning environment that combines qualified language proficiency and language teaching competency (Coniam & Faley, 2018).

CONCLUSION

Based on the data analysis, the researchers concluded that the results of the content validity analysis of the language proficiency test questions for English teachers are in line with the 'can-do' descriptors. Moreover, in the findings, the CEFR descriptors can be found in most LPATE test items. The LPATE test items also have covered almost all the language proficiency CEFR descriptors necessary to assess English teachers. This statement is also supported by one of the main benefits of CEFR framework which is to facilitate profiling the objectives for specific purposes.

Furthermore, the LPATE test can help to measure teachers' candidate competence to use the target language. This test can also be used to recruit and develop candidates with strong English communication skills. In the end, this test can be used to evaluate an English teacher's English speaking and comprehension skills.

In conclusion, since the contents of the English proficiency test for teachers or the LPATE test is valid and following the standard set by CEFR, then there is a necessity to evaluate the relation between the contents of the LPATE and the syllabus applied in Indonesia for future research on English language proficiency assessment for teachers.

REFERENCES

- Alderson, J.C. (2017). Foreword to the special issue "the common European framework of reference for languages (CEFR) for English language assessment in China" of language testing in Asia. *Language Testing in Asia*, 7(1), 1–9. <https://doi.org/10.1186/s40468-017-0049-9>
- Al-Malki, M.A.S. (2014). Testing the predictive validity of the IELTS test on Omani English candidates' professional competencies. *International Journal of Applied Linguistics and English Literature*, 3(5), 166–172. <https://doi.org/10.7575/aiac.ijalel.v.3n.5p.166>
- Andrew, S. (2007). *Teacher language awareness*. Cambridge: Cambridge University Press.
- Aniroh, K. (2009). From English as a general school subject onto English as a medium for learning specific subjects: The need to shift in the teaching orientation. *TEFLIN Journal*, 20(2), 169–179. <http://dx.doi.org/10.15639/teflinjournal.v20i2/169-179>
- Arikunto, S. (1992). *Prosedur penelitian*. Rineka Cipta.
- Arikunto, S. (2005). *Dasar-dasar evaluasi pendidikan*. Bumi Aksara.
- Aswal. (2020). *The analysis of the content validity of English formative test made by the English teacher at seventh grade of junior high school 2 Palopo*. Unpublished thesis.
- Bachman, L.F., & Palmer, S. (1996). *Language testing in practice*. Shanghai University Press.
- Brunfaut, T., & Harding, L. (2020). International language proficiency standards in the local context: Interpreting the CEFR in standard setting for exam reform in Luxembourg. *Assessment in Education: Principles, Policy & Practice*, 27(2), 215–231. <https://doi.org/10.1080/0969594X.2019.1700213>
- Coniam, D., & Falvey, P. (2018). *High-Stakes testing*. Springer.

- Correia, R.C. (2017). Assessing speaking proficiency: A challenge for the Portuguese EFL teacher. *e-Teals: An e-journal of Teacher Education and Applied Language Studies*, (7), 87–107. DOI:10.1515/eteals-2016-0009
- Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Dwiningtiyas, G.N., Sofyan, D., & Puspita, H. (2020). Teachers' strategies in teaching reading comprehension. *Journal of Applied Linguistic and Literacy*, 4(2), 66–77.
- Efset.org. CEFR dan EF SET. (2022). Retrieved from <https://www.efset.org/id/cefr/>
- Foley, J. (2019). Issues on the initial impact of CEFR in Thailand and the region. *Indonesian Journal of Applied Linguistics*, 9(2), 359–370. <https://doi.org/10.17509/ijal.v9i2.20233>
- Frost, J. (2022). *Content Validity: Definition, Examples & Measuring*. Retrieved from <https://statisticsbyjim.com/basics/content-validity/#:~:text=High%20content%20validity%20indicates%20that,knowledge%20of%20a%20subject%20area>
- Furwana, D. (2019). Validity and reliability of teacher-made English summative test at second grade of vocational high school 2 Palopo. *Language Circle: Journal of Language and Literature*, 13(2). <https://doi.org/10.15294/lc.v13i2.18967>
- Global Exam. (2019). *TOEFL vs IELTS: What are the differences between the tests?* Retrieved from <https://global-exam.com/blog/en/differences-ielts-toefl/>
- Heaton, J.B. (1998). *Writing English language test*. Longman.
- Hendryadi, H. (2017). Validitas isi: Tahap awal pengembangan kuesioner. *Jurnal Riset Manajemen dan Bisnis*, 2(2), 259334. DOI: 10.36226/jrmb.v2i2.47
- Huang, B.H., & Flores, B.B. (2018). The English language proficiency assessment for the 21st century (ELPA21). *Language Assessment Quarterly*, 15(4), 433–442. <https://doi.org/10.1080/15434303.2018.1549241>
- Hughes, A. (2003). *Testing for language teacher – second edition*. United Kingdom.
- Hughes, A., & Porter, D. (1983). *Current development in language testing*. London.
- Jin, D. (2020). Research on process evaluation for college Japanese teaching based on “Japan foundation can-do” standard. *Conference on Education, Language and Inter-cultural Communication (ELIC 2020)*, pp. 103–106. Atlantis Press. <https://doi.org/10.2991/assehr.k.201127.020>
- Kimura, Y., Nakata, Y., Ikeno, O., Naganuma, N., & Andrews, S. (2017). Developing classroom language assessment benchmarks for Japanese teachers of English as a foreign language. *Language Testing in Asia*, 7(1), 1–14.
- Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575.
- Little, D., & Bryam, M. (2004). *Encyclopedia of language teaching and learning*. Routledge.
- Nakata, Y. (2010). Improving the classroom language proficiency of non-native teachers of English: What and how? *RELC Journal*, 41(1), 76–90. doi: 10.1177/0033688210362617
- Nguyen, V.H., & Hamid, M.O. (2020). The CEFR as a national language policy in Vietnam: Insights from a sociogenetic analysis. *Journal of Multilingual and Multicultural Development*, 1(1), 1–13. <https://doi.org/10.1080/01434632.2020.1715416>
- Nugroho, H.M.A. (2018). EFL teachers' need of language proficiency professional development: When policy and practice collide. *International Journal of Language Education*, 2(2), 74–82. <https://doi.org/10.26858/ijole.v2i2.6373>
- Permendikbud Number 3. (2020). National Higher Education Standards.
- Pramana, C., Chamidah, D., Suyatno, S., Renadi, F., & Syaharuddin, S. (2021). Strategies to improved education quality in Indonesia: A review. *Turkish Online Journal of Qualitative Inquiry (TOJQI)*, 12(3), 1977–1994.
- Rachmawati, D.L., & Purwati, O. (2021). Linking language proficiency to teaching competence: A case study of EFL teachers in Indonesia. *Elsya: Journal of English Language Studies*, 3(3), 163–173. DOI: 10.1515/eteals-2016- 0009

Renandya, W.A., Hamied, F.A., & Joko, N. (2018). English language proficiency in Indonesia: Issues and prospects. *The Journal of Asia TEFL*, 15(3), 618–629. <https://doi.org/10.18823/asiatefl.2018.15.3.4.618>

Taylor, L., & Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalizing the test construct. *Journal of English for Academic Purposes*, 10(2), 89–101. <https://doi.org/10.1016/j.jeap.2011.03.002>

Thao, N. (2018). An Investigation into the content validity of a Vietnamese standardized test of English proficiency (Vstep.3-5) reading test. *VNU Journal of Foreign Studies*, 34(4). ISSN 2525–2445. <https://doi.org/10.25073/2525-2445/vnufs.4286>

TOEFL PBT®. (2022, December 9). *Paradigm Education*. Retrieved from <https://www.paradigm-edu.com/toefl-pbt/>

Tran, H.P., Griffin, P., & Nguyen, C. (2010). Validating the university entrance English test to the Vietnam National University: A conceptual framework and methodology. *Procedia-Social and Behavioral Sciences*, 2(2), 1295–1304. <https://doi.org/10.1016/j.sbspro.2010.03.190>

Uri, N.F.M., & Abd Aziz, M.S. (2018). Implementation of CEFR in Malaysia: Teachers’ awareness and the challenges. *3L: Language, Linguistics, Literature®*, 24(3). <http://dx.doi.org/10.17576/3L-2018-2403-13>.

Vanderbilt.edu. (2022). Retrieved from <https://www.vanderbilt.edu/elc/programs/english-academics-professionals/academic-speaking/>

William, W., & Jurs, S.J. (1990). *Educational measurement and testing* (2nd Ed., p.183). Boston.

Yan, J., & Huizhong, Y. (2006). The English proficiency of college and university students in China: As reflected in the CET. *Language, Culture and Curriculum*, 19(1), 21–36. DOI:10.1080/07908310608668752

APPENDIX

Reading Analysis

Reading Code	Item test
A1	
	1. What does the writer believe is special about the human species?
	2. Which of the following is the best paraphrase for the meaning of ‘To be sure’ (line 3) in this context? A. For us to be certain that B. Although it is certain that C. In order to be certain that D. We can be quite certain that
	3. According to the writer, why would a world full of solitary humans not seem remarkable to an extraterrestrial observer?
	4. What amazing fact about human is ‘captured in the story of the Tower of Babel’ (lines 5-6)
	5. How can having a common language help people to work in teams?
	6. What example of ‘ancient cooperation and shared ingenuity’ (line 16) does the writer give?
	7. What two words does the writer use in lines 16-19 to show that the conclusions are not completely certain?

	8. 'Chances are' (line 21) implies that: A. what follows is very likely to happen. B. what follows may happen by accident. C. people have opportunities to do what follows. D. there is a chance that what follows may happen.
	9. Why do people sometimes 'talk to themselves, to their dogs, even to their plants' (line 23)
	10. What are 'the spellbinding orator, the silver-tongued seducer, the persuasive child' (lines 24-25) examples of?
	11. How does the mention of aphasia in line 26 relate to the writer's main point in this paragraph?
	12. 'This book is about human language' (line 28). What is the writer contrasting human language with?
	13. '... there is something to write about it' (lines 33-34). What does 'it' refer to?
	14. '... in the years since' (line 37). In the years since what?
	15. '...these exciting discoveries' (line 40). What exciting discoveries?
	16. Which fields of inquiry are greatly affected by recent discoveries about the human language instinct?
	17. What 'event' is the writer referring to in line 46?
	18. According to most educated people, why do speakers of different languages 'construe reality in different ways' (line 48)?
	19. Which of the following are accurate reflections of the writer's opinion?
	20. 'It seemed like an epic blunder' (line 1). What does 'it' refer to?"
	21. Why did United Airlines decide not to go ahead with their announced changes?
	22. Which of the following is closest in meaning to 'pulled back the curtain on' (line 8)
	23. In the paragraph beginning on line 8, what example is there of 'gamelike techniques'?
	24. Why are workers expected to be more motivated when gamelike techniques are used?
	25. According to those who support 'gamelike motivational methods' (line 17), what should such methods not aim to achieve?
	26. Why can 'meaning' and 'excitement' be considered 'consolation prizes' (lines 18-19)
	27. Which of the following is closest in meaning to the phrase 'in line with' (line 21) as it is used in this text?
	28. What general conclusion can be drawn from the study published in 2012?
	29. Why does people's evaluation of probabilities (line 30) make lotteries in general more effective than fixed payments?
	30. What example does the writer give of 'simple compliments and recognition' (line 32)?

	31. In the study published in 2014, how was the effectiveness of the experimental strategy measured?
	32. Which of the following best summarises the content of this passage?
	33. 'It's not the first book...' (line 2). What does 'It' refer to?
	34. Who is the 'moustachioed French philosopher' (line 5)?
	35. What does the phrase 'detonates a truth bomb' (line 5) suggest about the main idea to be found in <i>Time and Free Will</i> ?
	36. What phrase in Henri Bergson's is the writer paraphrasing when he refers to 'all the wonderful things you could do' (lines 10-11).?
	37. What do we have to give up when we make a decision?
	38. In the paragraph beginning on line 16, what phrase does the writer use to indicate that reading <i>Time and Free Will</i> fundamentally changed his ideas about decision-making?
	39. According to the writer, what is the implication of the fact that 'you only get the one life' (line 19)?
	40. Why does the writer feel 'almost inexpressible relief' (line 23)?
	41. What are the two errors that older people tend to make when they romanticise their youth?
	42. Explain 'most of our dreams were guaranteed to bite the dust' (lines 27-28)
	43. Why is it incorrect to believe that if you had your time again you would make better decisions?
	44. What is 'uplifting message' (line 36) for middle-aged people?
	45. What is one advantage of being old?
	46. Which THREE of the following statements are accurate reflections of what the writer says in this passage?

Writing Skill

Writing Code	Item test
A2	<p>No. 47</p> <p>Task 1</p> <p>As part of a Professional Development Day at your school, you have been asked to give a short talk to other teachers on the importance of students developing healthy friendships and avoiding ‘toxic friendships’.</p> <p>Write a short talk explaining why healthy friendships are important for successful academic achievement and how teachers can help students get out of ‘toxic friendships’.</p> <p>Write about 400 words.</p>
A3	<p>No. 48</p> <p>Part 2</p> <p>Task 2A: Detection and correction of errors/problems</p> <p>Below is a student’s composition. Although it contains a large number of errors/problems, you are asked to correct those in Questions 2 to 11 only.</p> <p>Each question has <i>at least one error/problem and possibly more</i> (e.g., subject-verb agreement or omission of third person singular). When you provide correction, please retain the original meaning and words as much possible; any unnecessary changes may be penalised. Questions 1 has been done for you as an example in the right-hand column.</p>
	<p>No. 49</p> <p>Task 2B: Explanation of errors/problems</p> <p>In this task, you are asked to complete the explanations of the errors/problems in Questions 12 to 20 to show your understanding of them. You may use more than one word to fill each blank.</p> <p>You should demonstrate to the examiners your understanding of the underlying rules or generalisations, using grammatical terms where appropriate to complete the explanations.</p>

Listening Skill

Listening Code	Item test
A4	<p>Part 1.</p> <p>You are going to listen to a radio show about fortune cookies. Hosts, Marco Werman and Lydia Kim, interview two fortune writers, James Wong and Russell Rowland.</p>
	50. According to Lydia, which of the following is typical associated with Chinese American meals?
	51. How did fortune cookies end up being served in restaurants in America? Complete the note sheet.
	52. Which of the following matches the meaning of the Japanese proverb?
	53. Why was James Wong offered the job of fortune writer at Wonton Foods?
	54. According to James, what is the difference between a fortune teller and fortune writer?
	55. Where does James get his inspiration to write his fortunes?
	56. What is the biggest challenge for James as a fortune writer?
	57. What guidelines does James offer for writing fortunes?
	58. Complete the story about a fortune that went wrong.
	59. Why did fortune cookies make the news on November 4, 2000?
	60. What did Russell Rowland think about the job interview as a fortune writer?
	61. What does Lydia think about the amount Russell Rowland was paid for writing fortunes?
	62. Why did Russell need the distraction of writing fortunes?
	63. Where were the ideas for some of Russell's fortunes taken from?
	64. How did Russell finish writing the last 200 fortunes?
A5	<p>Part 2</p> <p>The Secret Strength of Introverts</p> <p>In this podcast, Susan Cain discusses introversion and shares tips of teaching quiet kids in the classroom. Guests include Lara, a parent, Dr. Schwartz, a researcher and Jess, a teacher.</p>
	65. What is the aim of the Quite Revolution?

	66. Name TWO positive attributes of introverts mentioned by Susan Cain
	67. What are three signs that you could be an introvert?
	68. Why didn't Lara's daughter, Margaret, like the Taylor Swift concert?
	69. What does Susan recall not liking as a child?
	70. Complete the information about Dr. Schwartz's research.
	71. After hearing the comment made by her teacher, why did Susan feel upset?
	72. Name two limitations about class participation that Jess identifies.
	73. What strategies did quiet students adopt when participating in class?
	74. What criticism does Susan make about the education system?
A6	Part 3 Graduation Speech
	You are going to listen to a speech delivered by guest speaker Maria Popova at a graduation ceremony. She offers some life advice to graduates of the University of Pennsylvania.
	75. Complete the profile of Maria Popova.
	76. What stereotype does Maria mention about New Yorkers?
	77. Maria sees spring as a time when people_____.
	78. Listen to Maria's bicycle story. Put the events below in the correct order. Write 1-6 in the boxes provided.
	79. Reflecting on the bicycle incident, why did Maria feel lucky? When she was a student, why did Maria feel like everyone around her was on an electric bike?
	80. When comparing herself to others, what did Maria feel insecure about and why?
	81. Complete the note-sheet which outlines the progressive stages of soul's destruction.
	82. What is 'intelligent hope'? Tick ALL that apply.
	83. Referring to Maria's advice for the new graduates, complete the tab below.
	84. According to maria, what is badly needed today?

Speaking Skill

Speaking Code	Item test
A7	<p>NO. 85</p> <p>Task 1A Reading Aloud</p> <p>Please read the following prose passage aloud as meaningfully as possible.</p>
A8	<p>NO. 86</p> <p>Task 1B Recounting an Experience / Presenting an Argument</p> <p>The what extent is it beneficial to society to engage retirees in life-long learning? Give reasons for your answer.</p>
A9	<p>No. 87</p> <p>Task:</p> <p>Discuss collaborative learning in the English Classroom.</p>

Classroom Language Assessment

Classroom Language Assessment Code	Item test
A10	<p>88. Grammatical and Lexical Accuracy and range</p> <p>This scale asks the candidates to display the ability to communicate effectively with appropriate use of grammar. They also must demonstrate accurate use of an extensive range of grammatical structures, and they also have a sufficient vocabulary to make meaning clear.</p>
A11	<p>89. Pronunciation, Stress, and Intonation</p> <p>The candidates must be able to have an accurate pronunciation, stress and intonation. The good performance is characterised by smooth and natural flow speech enable by effective use of linking features and fine variations.</p>
A12	<p>90. Language of Interaction</p> <p>The candidates must be able to properly phrase a variety of questions, including open-ended questions, to encourage opinions, elaboration and discussion, as well as to give feedback.</p>
A13	<p>91. Language of Instruction</p> <p>The candidate must have an ability to spontaneously adjust the use of functional language when needed.</p>