

Dr .Mohammad Almutairi

Reliability and Validity of the Electronic English Placement Test (EPT) conducted by the Language Centre, Public Authority for Applied Education & Training

Dr .Mohammad Almutairi

College of Basic Education,

Language Center,

Public Authority for Applied Education and Training, Kuwait.

Abstract:

This study aims to investigate the validity and reliability of the electronic English language test administered by the Language Centre, the Public Authority for Applied Education and Training. 10 EFL teachers were interviewed by the researcher using a semi-structured interview. 100 EFL freshmen students were asked to perform the EPT for the second time and then some of them participated in a semi-structured interview as well. The researcher used the mixed method design in his study, the quantitative and qualitative methods, which added strength to the data analysis. The findings showed that EPT had deficient levels of validity and reliability. The reasons behind this result varied from ignoring writing, listening, and speaking when designing the EPT to poor instruction procedures and proctoring. The lack of new and advanced computers did affect the validity and reliability in a significant way. This study ends with recommendations to improve EPT's validity and reliability levels.

Keywords: ELT, TESOL, Testing, Assessment, Language Education

Reliability and Validity of the Electronic English Placement Test (EPT) conducted by the Language Centre, Public Authority for Applied Education & Training

Dr .Mohammad Almutairi

College of Basic Education,
Language Center,

Public Authority for Applied Education and Training, Kuwait.

Introduction

The Public Authority for Applied Education and Training (PAAET) is a higher-educational institution that consists of five colleges: the College of Business Studies, the College of Basic Education, the College of Medical Science, the College of Nursing, and the College of Technological Studies. All the mentioned colleges require applicants to complete the EPT, which is a test designed for high school graduates who plan to continue their education in one of the colleges. The EPT is conducted by the Language Centre twice per year, once at the beginning of each semester. All accepted students in the five colleges must take the exam before the beginning of the academic year. The test consists of 100 questions, with 60/100 serving as a passing score. If a student passes the exam, he or she can register in English 1 and other levels of English, but if the student doesn't pass, he or she has to join the remedial course, English 099.

Since being transferred by Teams Software to an electronic method of implementation, this test has raised lots of discussions among EFL teachers. Some of them think it has many disadvantages and should be improved to accomplish its main goal, which is to place students at their actual language proficiency level. Some, alternatively, think the EPT suffers from severe drawbacks and weaknesses and should be reconstructed.

Purpose of the study

The purpose of this study is to investigate the validity and reliability of the EPT conducted in the Language Centre at the PAAET, which is designed for high school graduates, and to offer the best suggestions and ways to improve it. This study is expected to improve English language learning and teaching in the PAAET colleges by improving the EPT, which, in turn, will help place EFL learners at their actual EFL levels.

Research questions

- 1- How valid is the EPT according to the views and opinions of EFL teachers in the Language Center at PAAET?
- 2- How reliable is the EPT according to the views and opinions of EFL teachers in the Language Center at PAAET?
- 3- What are EFL teachers' views and opinions towards improving the current EPT?

Importance of the English placement test

Many higher institutions tend to design language placement tests to identify students' levels of language proficiency to be able to place them in the right learning levels (Lamb 2017). When supported with test results, EFL students and EFL teachers are able to decide on the course that is most appropriate for them. A higher-level language class not supported by test results wouldn't suit the student's learning ability, while a lower-level class would be discouraging and reduce language-learning motivations. The EPT results assist in selecting the appropriate English course.

Context of the study

This study takes place in the Language Centre, an institution that undergoes the Public Authority for Applied Education and Training. PAAET was established on December 28, 1982, by the Royal Amiri Decree number (63) to develop and upgrade manpower to meet the challenge of a shortfall in technical manpower created by the industrial and economic development of the country (Nawfl & Alhindi, 2001). The foundation of applied education and training was laid along with the initiation of oil exploration, production, and

export in Kuwait. In the 1950s, the State began to establish training centers and organized programs to prepare the manpower needed for the oil industry. After building the fundamental structure of the educational system, the Ministry of Education established several specialized institutions to meet the increasing demand for skilled manpower. The other ministries established their training centres and institutes as well. The State then found it essential to establish a central body to supervise and coordinate the activities of these numerous institutes.

The Technical and Vocational Education Department was established in 1972 to supervise technical and vocational education. The Central Training Department was established to supervise training centers and institutes set up by State Ministries.

The Technical and Vocational Education Department continued to function till 1982 when the Public Authority for Applied Education and Training (PAAET) was established as an autonomous body to supervise technical and vocational training and envision and implement an overall plan for the sector.

The goal of PAAET is to develop national technical manpower and to meet the human resource needs of the country through its two sectors: Education and Training.

The applied education sector includes 5 colleges, which offer several specializations: The College of Basic Education, The College of Business Studies, The College of Technological Studies, The College of Nursing and The College of Health Sciences.

The main goal of the College of Basic Education which awards the degree of (Bachelor in Education) is to provide government schools with qualified Kuwaiti teachers for the kindergarten, primary and intermediate stages. The duration of its programs is four years or eight semesters.

The main goal of the College of Business Studies which awards only diploma certificates is to provide the two sectors of the market – the government and the private sector – with Kuwaiti employees specialized in banking, finance, insurance, and management. The

duration of the study at the CBS is from two-and-a-half years to three years.

The College of Technological Studies awards only diploma certificates in mechanical, civil, electrical, computer, and petroleum engineering. Its goal is to provide the labor market with Kuwaiti assistant engineers in such fields.

The College of Nursing and the College of Health Studies award diploma certificates in nursing, radiology, medical lab technology, and nutrition, their goal is to provide the Kuwait Ministry of Health with Kuwaiti nursing staff and other related medical specialties due to the shortage of national workers in such fields.

Besides the four colleges, there is the Language Centre which contains, in addition to the headquarter, several language units distributed among all the PAAET colleges. The main objective of the language units is to teach the PAAET students the general English they need in addition to ESP according to their specialties.

Methodology

As Cohen, Manion, & Morrison (2007) have argued, the purpose of the research determines the research methods. As mentioned earlier, the purpose of this study is to test the validity and reliability of the English language placement test conducted by PAAET's Language Centre. In order to gather the required data for the study, the researcher decided to choose the mixed method design and use both quantitative and qualitative methods. As for the quantitative methods, 135 students were asked to perform a language test consisting of 100 questions. 100 students agreed to participate. For the qualitative method, a semi-structured interview with 5 EFL teachers and 10 EFL students only was used.

Some researchers like Johnston (2000) think that a quantitative approach that includes questionnaires introduces the research questions in a way in which the results or outcomes can be measured in numbers, calculations and statistics, that is, in a way that searches for the answer of how often something or a phenomenon happen.

The qualitative method is considered one of the best methods in educational research as it addresses the “how” and “why” research questions and allows a deeper understanding of experiences, phenomena, and context (Cleland 2017). The qualitative method permits scholars and researchers to ask questions that cannot be easily presented in numbers in order to investigate a significant issue.

Moreover, the qualitative approach, including semi-structured interviews, introduces the research questions in interpretative ways, such as asking questions about watching or listening to the phenomena under investigation – that is, searching for the answers to why and how things happen.

The semi-structured interviews are chosen and prepared carefully in this study to explore the topic in more depth and because they allow the interviewees to express their opinions, concerns, and feelings in a comfortable way (Shohamy & Seliger 1989). They have many advantages that encourage researchers, especially in education, to use them. The major advantage of semi-structured interviews is the confidential atmosphere in which participants can share sensitive information and details about their personal experiences, views, and behaviour. A private setting also means that peers do not influence participants’ responses to study questions (Hogle & Sweat 1996).

Benefits of mixing methods

As mentioned before, the researcher uses both quantitative and qualitative methods to gather the data needed for the study. Several researchers, such as Creswell (1994), have argued that by combining the methods, researchers can have a better understanding of the phenomenon under study and can add breadth and scope to the study. Moreover, as Frechtling, Sharp, & Westat (1997) argued, mixing the methods (quantitative and qualitative) can increase the validity and reliability of the study and sharpen our understanding of research findings: Data from interviews can clarify data collected from a questionnaire. Also, they indicated that a mixed-method design in a single study can strengthen the validity of the research results.

Merriam (1988) also argued that quantitative data from questionnaires or other instruments can be utilised to support findings from qualitative data.

Research Tools

As mentioned earlier, the researcher used 2 semi- structured interviewers for EFL teachers and EFL learners. The semi- structured interview designed for EFL teachers included the following questions:

- Q1- How valid is the English Placement Test from your opinion?
- Q2- How reliable is the English Placement Test from your opinion?
- Q3- What are the advantages of the English Placement Test if available?
- Q4- What are the drawbacks of the English Placement Test if available?
- Q5-What are your suggestions for improving the English Placement Test?

The semi- structured – interview designed for EFL learners included the following questions:

- Q1- Can you describe your experience during conducting your English Placement Test?
- Q2- What are your views about the validity or the reliability of the English Placement Test?
- Q3- What are your advices to improve the English Placement Test?

Literature Review

Most Arabic students, including Kuwaiti, face significant educational and linguistic hindrances due to the transition from Arabic medium to English at the university level. As a result of this transition, the first year is usually designed to equip high school graduates to succeed in their English medium instruction studies and become fluent communicators (Kadwa & Sheik 2021).

The first step toward offering the new students suitable and successful language learning is placing them in a language placement test to identify their language levels. Depending on their

scores on the placement test, EFL students could be distributed among the appropriate language levels.

In other words, placement tests are used by higher institutions to determine student accomplishment levels and then place them in the appropriate courses based on their current skill level (Barnett 2017). According to Heaton (1988) and Brown (2004), placement tests are used to assess the students' language and linguistic ability in the target language to direct them to the appropriate level of language learning. Accurate and effective test assessment is essential for successful foreign language learning.

In recent decades, the effectiveness of placement tests has become very controversial among researchers. As a result of that debate, two main elements came to the surface that should be considered when planning any placement test: validity and reliability.

Today's perspective on validity in language testing is derived from an article by Messick (1989), which has influenced educational and psychological measurement as well as language testing. Even though pre-Messick concepts, expressions, and perspectives continue to exist, particularly in language testing, Messick's presentation of validity is widely accepted in language testing today. According to Brown and Abeywickrama (2010) and Lado (1961), validity is considered a trait of tests. According to them, a test becomes valid if it measures what it has to measure and nothing more.

The importance of a valid and a reliable test is in placing students into the appropriate levels of learning. Conversely, an invalid or unreliable test leads students to the wrong learning levels, which might have a negative impact on their proficiency and could also lead them to develop bad attitudes towards language learning (Al-Adawi & Al-Balushi 2016). In addition, the teaching and learning processes can be a burden for both teachers and students when students are at the wrong language level (Johnson & Riazi 2017). According to Hille and Cho (2020), invalid or unreliable placement

tests may lead to financial implications, affect students' degree plans, and reduce their motivation levels.

As for reliability, it is one of the most important characteristics of all tests in general, and language tests in particular. The reliability of a test concerns the consistency of scoring and the accuracy of the administration procedures of the test (Chiedu & Omenogor 2014).

Najib (1999) explains that reliability refers to the consistency of test results. He adds that the reliability level becomes high if a student has a certain skill level and demonstrates the same level for a second time. Reliability, to him, can be determined by the test-retest.

For Drost (2011), reliability is "The extent to which measurements are repeatable when different people perform the measurement on a different occasion, under different conditions, supposedly with alternative instruments which measure the construct or skill".

Results and discussion

As this study aims to identify the validity and reliability of the English placement test conducted by the Language Center in the PAAET, the researcher started interviewing some EFL teachers who designed the test, conduct it twice a year, do the proctoring each semester, and teach the examined students their EFL classes.

Also, a hundred EFL students who passed the previous English placement test with scores of 60/100 and higher were asked to perform the test for a second time a week after conducting the first one.

After conducting the test for the second time and marking their answers, the researcher started analysing the data, and the students' results were put in a table for comparison (Table 1). The findings were unexpected, as sixty students out of a hundred failed the exam and got below 60/100 and forty students passed the exam in various grades. These findings obviously show that the reliability of the English measurement test is very low. This high percentage of failure on the test certainly occurred for reasons.

To discover the reasons behind these astonishing test results, the researcher started interviewing the failing students. Most of them

admitted that they got some help while performing the first test, and some of them revealed that they got help from other students, benefiting from poor proctoring. Student A, for example, said, “I knew earlier from my old colleagues, who preceded me in the college that I could use my mobile to get answers and to get help.” Many students, according to him, received pictures of the test questions with the correct answers on their mobiles while the test was being conducted. Moreover, student B said that he got a high grade on the first test through outside help, as he took many answers from the students sitting next to him because the desks in the computer labs were stuck to each other. Student C, on the other hand, claimed that he had passed the previous test because the test was multiple choice, which encouraged him to answer randomly the questions he could not answer.

Student D, who also failed to pass the EPT on the second attempt, expressed his sympathy for his colleagues, as they weren’t given enough time to answer their questions.

He clarified that the test contained 100 questions and that each question had four multiple-choice answers. Students were given only 55 minutes to read, think, answer, revise, and submit the exam. To him, the reliability of the test was not even acceptable because the atmosphere of the EPT on the boys’ campus was not encouraging. “It was a complete mess,” he said. He clarified this important point by saying, “The number of the students is around 6,800 who have to be tested within two days. The crowded corridors and the crowded computer labs did not offer the students the relaxing atmosphere that they needed to concentrate on the questions.” He also added that a lot of cheating took place in the previous exam because students were allowed to use their mobile phones and because of poor proctoring.

When interviewing the passing students about their opinions and views about the validity and reliability of the EPT, most of them said that the test was very low in its validity and reliability for certain reasons.

They believed that a valid test should include questions about topics and grammar that were studied in the high school syllabus. For them, some of the content of the test in the reading passages and grammar was new.

Student G, for example, said, “Although the grammar tenses in the test were part of the syllabus in high school, the questions of the grammar contained a lot of new vocabulary.” According to him, a valid test shouldn’t contain new topics or new vocabulary.

He ended his opinion by saying that the English placement test in PAAET should be designed carefully from the high school syllabus if PAAET wanted to measure the new freshmen students.

	Previous score range	New score	Result 1 + result 2	Percentage of change
10 Students	60-64	54-58	Pass – Fail	28%
10 Students	65-69	50-54	Pass – Fail	35.9%
10 Students	70-74	39-51	Pass – Fail	51.3%
10 Students	75-79	41-58	Pass – Fail	50%
10 Students	80-84	35-52	Pass – Fail	57.7%
10 Students	85-88	32-56	Pass – Fail	58%
10 Students	89-90	67-71	Pass – Pass	13%
10 Students	91-93	69-75	Pass – Pass	17%
10 Students	95-97	72-81	Pass – Pass	10%
10 Students	98-100	74-83	Pass – Pass	6.4%

(Table 1)

As mentioned earlier, to get answers to the first two research questions, the researcher interviewed 10 EFL teachers in their offices in the Language Centre.

The analysis of EFL teachers’ responses by the semi-structured interviews for Question 1 (“How valid is the EPT according to the views and opinions of EFL teachers in the Language Center in PAAET?”) showed that ten teachers out of ten (100%) believed that the English placement test was not valid for some reason.

Teacher A, for example, who has PhD in applied linguistics from the UK, stated that a valid English language test should include testing the four language skills: listening, speaking, reading, and writing. Unfortunately, the current English language tests, according to him, test only reading and grammar, which doesn't give the actual students' language proficiency.

He also clarifies why it is very difficult to test students' proficiency in listening. He claims that the shortage of technical facilities, such as language labs, needed for listening tests in the Language Centre makes it impossible to conduct such listening skills testing. He also adds that the shortage of teachers and the need to deliver the results to the Registration Office within 48 hours make it very difficult to conduct speaking skills tests for students.

He ensures that increasing the validity of the current English placement test would come through the inclusion of speaking, listening, and writing in the next EPT.

In his answers to the second research question, the reliability of the EPT, Teacher A thinks that the accuracy of the administration procedures of the test is part of the reliability of any test. According to him, test designers should design their test questions using clear instructions and directions. He states that tested students do not receive any instructions or advice during the performance of their tests. He adds that this is clear evidence of the low reliability of the current EPT.

Teacher B, who has an MA in TESOL from the USA and who has 14 years of experience in teaching English as a foreign language in PAAET, also shares Teacher A's opinion that the validity of the current English placement test is very low. He thinks testing vocabulary and grammar only does not truly give the real language level of proficiency and consequently places EFL students in the wrong English classes. He also understands that testing listening skills requires special technical devices that are not available for the Language Centre, but he questions why testing writing skills is missing. He adds that testing writing skills should have been

included in the past English placement tests because it requires no technical facilities. In his opinion, the current EPT suffers from great challenges and should be reconstructed. To increase its validity, he suggests the missing skills should be included in the next test design.

Regarding the reliability of the EPT, he explores a great challenge that affects the reliability each time the test is administered: the way of conducting the exam. He adds that many male students face a lot of technical problems with their computers in the computer labs because they are old and consumed. Sometimes the Teams software hangs for more than 10 minutes or submits the test early. On other occasions, Teams doesn't start on time for students. In all cases, the student has 55 minutes to finish his or her test and one attempt to perform the test. This, according to Teacher B, is a great drawback to the test that decreases its reliability. As a solution to this problem, he suggests that the PAAET should provide the labs with new computers.

Teacher C, who has an MA in applied linguistics and has 11 years of experience in teaching English as a foreign language, explores the idea that the test should be reorganized. First, she suggests that the test should be rewritten to include the four language skills and be divided into certain periods. She posits that the testing of reading and writing skills should occur on a different day than the testing of speaking and listening skills. This, according to her, would help to ascertain the real language proficiency level of test takers, which, as a result, would increase the validity of the EPT. She adds that the current EPT consists of 100 multiple-choice questions, which does not help in accurately determining students' language levels because students could pass the EPT via random guessing. With regard to the reliability of the test, she believes that the more time students are given to complete the test, the more valid and reliable the test's results get. She thinks that giving students 55 minutes to answer 100 questions is not sufficient. She believes that the circumstances the students experience during the test, such as tension and pressure,

make the 55 minutes insufficient for answering the long queue of questions.

Limitations of the study

The researcher believes that it very important to mention the limitations of his study. The small sample size is a clear limitation which should be emphasized in any future research. Because the short number of participants could have led to an overestimation of the status of the English Placement Test in this study, future research should reconfirm these findings by conducting larger-scale studies.

Conclusion and recommendations

The purpose of this study was to uncover the validity and reliability of the English placement test at the PAAET. The findings of this study showed that the current EPT is invalid and unreliable for several reasons. The researcher thinks that improving the test's validity could be achieved by including the four language skills – listening, speaking, reading, and writing – in the test. Extending the period of conducting the test up to 5 working days would help EFL teachers manage and enhance the instruction procedures. Extra time would also enable EFL teachers to do true proctoring, which, in turn, prevents the tested students from securing any means of external help. Increasing the amount of time that the students have to complete the test would improve the validity and reliability of the EPT, as the longer the test's duration is, the more reliable and valid it becomes. Furthermore, in practice, reliability is increased by making the test's instructions very clear, reducing the scope for variety in the answers, and making sure the test items stay constant.

References

- Al-Adaw S. et al (2016). Investigating Content and Face Validity of English Language Placement Test Designed by Colleges of Applied Sciences. *English Language Teaching*; Vol. 9, No. 1; 2016.
- Barnett, E. A., & Reddy, V. (2017). *College Placement Strategies Evolving Considerations and Practices. A CAPR Working Paper*. Retrieved from <https://files.eric.ed.gov/fulltext/ED583509.pdf>.
- Brown, H.D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, H. D. & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practice* (2nd ed.). New York: Pearson Education Inc.
- Chiedu, R. E., & Omenogor, H. D. (2014). The Concept of Reliability in Language Testing: Issues and Solutions. *Journal of Resourcefulness and Distinction*, 8(1), 1-9.
- Cleland, J.A. (2017). The Qualitative Orientation in Medical Education Research. *Korean J Med Educ*. 2017 Jun; 29(2): 61–71.
- Cohen, L. et al (2007). *Research Methods in Education* (5th ed.). London: Routledge.
- Cresswell, W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, Calif: Sage Publications.
- Drost E. (2011). Validity and Reliability in Social Science Research. *International Perspectives on Higher Education Research* 38(1):105-124.
- Heaton, J.B. (1988). *Writing English Language Tests*. Harlow: Longman.
- Hille, Kathryn & Yeonsuk Cho. 2020. Placement testing: One test, two tests, three tests? How many tests are sufficient? *Language Testing* 37(3). 453–471.
- Hogle, J. and M. Sweat. (1996). *FHI/AIDSCAP Evaluation Tools: Qualitative Methods for Evaluation Research in HIV/AIDS Prevention Programming*. Chapel Hill, NC: Family Health International.
- Johnson R. C., Riazi A. M. (2017). Validation of a locally created and rated writing test used for placement in a higher education

- EFL program. *Assessing Writing*, 32, 85–104. <https://doi.org/10.1016/j.asw.2016.09.002>
- Johnston, B. (2000). Investigating dialogue in language teacher education: The teacher educator as learner. In K. Johnson (Ed.), *Teacher education* (p. 157-173). Alexandria, VA: TESOL.
- Kadwa, M. S., & Sheik, A. (2021). A Critical Review of the Oxford Q-Skills Placement Test at a Saudi Arabian University. *International Journal of Linguistics, Literature and Translation*, 4 (1), 198-209.
- Lamb, M. (2017). The motivational dimension of language teaching. *Language Teaching*, 50 (3), 301-346.
- Lado, R. (1961). *Language testing*. London: Longmans.
- Merriam, S.B. (1988) : *Case study research in education*. 2nd ed. Jossey-Bass Inc. Publishers, San Francisco.
- Najib, M. (1999). *Research in Education*. University of Technology, Malaysia.
- Nawfl, E. and Alhindi, K. (2001). *Mkhrjat alta'aleem wa hajat souk al'amal fi al kuwait* [Educational outcomes and labour market needs in Kuwait]. Kuwait: Ministry of Education, Department of Educational Research.
- Seliger, H. & Shohamy, E. (1989), *Second Language Research Methods*, Oxford: OUP.