

DOI: <https://doi.org/10.38035/dijemss.v5i6>

Received: 16 August 2024, Revised: 20 August 2024, Publish: 2 September 2024

<https://creativecommons.org/licenses/by/4.0/>

An Analysis on Final Term Test of Pre-Intermediate Level in Healthcare English Program for Midwifery

Sarovah Widiawati¹, R.M. Rizky Wirabrata², Novi Indah Susanthi³, Riza Lestari⁴¹ITL Trisakti, Jakarta, Indonesia, sarovahwidiawati@gmail.com²National English Centre, Jakarta, Indonesia, rizky.wirabrata@gmail.com³ITL Trisakti, Jakarta, Indonesia, oct4th2003@gmail.com⁴ITL Trisakti, Jakarta, Indonesia, rizalestari@gmail.com*Corresponding Author: sarovahwidiawati@gmail.com

Abstract: This research aims to find out to what extent that final-term test of pre-intermediate level in healthcare English program complies to the criteria of good test. A qualitative design with content analysis method was carried out to conduct this study. To analyze the data, the researcher first examined the test items and instructions in both the oral test and final test. After that, the items were classified. Once it was done, the researcher would decide to what extent the test is considered as a good test. The result shows that the test does not fulfill the criteria of a good test since only some of the criteria are complied. The oral test contains only three principles, which are validity, authenticity, and washback. Meanwhile, the same goes to the written test since it only covers validity, authenticity, and washback.

Keywords: Analysis, final-term test of pre-intermediate level, test, principles, criteria of a good test.

INTRODUCTION

In a course, achieving the goal of the learning has to be the main focus since it pinpoints the standard on what the students should be able to do at the end of the program. In order to ensure a goal has been fulfilled, the teacher ought to measure the students' knowledge before the course ends. That activity is called an assessment, which is widely considered as one of the most important aspects in teaching and learning activity. Should it be overlooked, teachers will hardly identify if the learning outcomes have been achieved. In addition, in order to carry out an assessment, a test, as the instrument, is required.

A test has to be meticulously designed by teachers in order to measure the learners' knowledge properly. To fulfill that, the test developer should find a reference from the experts. Brown (2004) proposes criteria for a good test that encompasses practicality, validity, authenticity, reliability, and washback. Thus, this draws the author attention to conduct a study to analyze a test.

The type of test analyzed in this research is the achievement test. Achievement test is related to classroom lessons and units related to the curriculum at the end of teaching and learning activities (Brown, 2004). The test is the final term test for the pre-intermediate level. It is provided for midwifery classes in several midwifery academies and universities that

collaborate with National English Centre. It is a kind of English for Specific Purposes test where English is a medium in supporting teaching and learning activities in facing the globalization era. Farnaz et al (Latif, 2013) stated ESP test has several features to enhance students' communicative performance, to reflect on the learning process, and to encourage learning, and supervise the learner's progress.

Table 1. Final test outline, pre intermediate level

Oral test Presenting a procedure (how to bathe the baby, how to give intramuscular injection, and how to do vulva hygiene)
Written Test Part 1. Listening (10 items) : Multiple choice Part 2. Reading (10 items) : Multiple choice Part 3. Language focus (10 items): Circle the correct answer Part 4. Writing (10 items). Match definitions and procedures

The final test consists of two parts, oral test and written test. The former has the students present a medical care to a general practitioner/senior midwife who has responsibility in that field as the examiner while the latter is a multiple-choice question test. In the oral test, there are medical instruments to support the presentation, like syringes, gloves, alcohol swabs, liquid medicine, and kidney dishes. Moreover, the students are divided into a group of four; each group has different topic. The test lasts for 20 minutes for each group. In addition, the written test consists of 40 questions, 10 questions for listening, 10 questions for reading, 10 questions for language focus (grammar), and 10 questions for writing. Five major points were used to analyze the test; they are practicality, validity, reliability, authenticity, and washback. Furthermore, the syllabus and student books were employed to support the analysis. The purpose of this research is to identify to what extent that the final-term test of Pre-Intermediate Level fulfills the criteria of a good test proposed by Brown (2004).

A test is an essential part of teaching and learning activities (Heaton, 1989). Brown stated that it is a technique used to measure students' ability, knowledge, or performance in a certain field (Brown, 2004). A test is used by a teacher to identify the students' progress after teaching and learning activities. In addition, it also serves as an evaluation tool in controlling and determining the quality of teaching and learning activities (Elthia & Alawiyah, 2023).

Tests may vary in types depending on the purposes of the test itself. Several kinds of tests are multiple-choice, essay, cloze test, matching, oral, interview, and checklist. Oral test has advantages for students. Mahendra (2023) stated that the students who took oral tests felt worried about the exam, but they considered that the test gave more advantages than other forms of test. There are five types of oral tests, they are imitative, intensive, responsive, interactive, and extensive (Ahsanu, 2013). Moreover, speeches, oral presentations, and storytelling are categorized as extensive oral tests. The use of a checklist or grid is also common for scoring the test. To be qualified as a good test, the test should meet the criteria of being one. Brown stated that there are three criteria of a good test, practicality, validity, and reliability (Brown, 2001). Practicality means the test should be arranged with a low budget, time limitation, implementation, and scoring system. Validity is the ability of the test to deal with learning goals and competencies that should be assessed while reliability refers to the consistency of the test.

METHOD

This research used qualitative design with content analysis method. Cresswell (1994) described qualitative research as an unfolding model that occurs in a natural setting that enables the researcher to develop a level of detail from high involvement in the actual experiences. search involves describing, explaining, and interpreting collected. Furthermore, this research

also employed content-analysis method. Fraenkel and Wallen (2009) argues that content analysis refers to a method that allows the researcher to learn people's communication indirectly. Generally, content analysis is largely used to analyze written documents, such as: textbooks, essays, newspapers, test items, and many more.

This research used the final test of Pre-Intermediate English for midwifery document as the object of the test. Considering that this is a qualitative research, the researcher is the main instrument (Fraenkel & Wallen, 2009). The data is the test items in the final test of Pre-Intermediate English for midwifery. There are 40 questions, and all of them are multiple-choice items. To analyze the data, the researcher began with checking each item in the test, including the stem and the options. Next, the item was classified whether it was considered as a good test item or not based on the 5 criteria by Brown (2004). Finally, after all items were classified, the researcher decided if the whole test had all the criteria.

RESULT AND DISCUSSION

Practicality

Practicality means effective. It has four criteria (Brown, 2004).

- a. The cost is affordable.
- b. There is time limit.
- c. The test administration does not cause burdens.
- d. The test is easy to score, in terms of efficiency and effectiveness.

Based on the criteria, the oral test is impractical because it needs a general practitioner or senior midwife to conduct the test. Moreover, the general practitioner is not involved in the teaching and learning activities, so they may have biased scoring. The duration of the test was 30 minutes for each group because each member of the group should present the procedure. One group consists of four to five students. There were more than 5 groups in the test. While the other groups represent the oral test, the other groups should wait their turn in more than 10 minutes. In administration, the test has moderate practicality because the test should be held in the laboratory which provides the instruments based on the topic that they get. Where the laboratory in each academic and university is not in good condition, the air conditioner or the cleanness is terrible, and the instruments are limited, like phantom, syringe, and wash basin. Because of that the students preferred to bring the instruments into their class and set the class into a laboratory. The test also should be recorded as documentation; here the facilitator should record and organize the oral test. There was no score list standard to measure the oral test. There are two different interpretations in assessing the oral test, like a general practitioner focuses on students' fluency and knowledge and a senior midwife focuses on fluency, confidence, and knowledge.

On the other hand, the written test is highly practical. It is because the test is not expensive. It is only two pages in one paper of A4 so the cost for photocopy is cheaper. Each student gets the problem sheet and they could answer that problem on the problem sheet directly without the answer sheet. The listening test is also already provided by the institution in one folder with the label of each level and test, so it makes the proctor play it, but it is impractical because not all of the universities or academics provide audio for listening. The time allocation for writing tests is flexible. It is two hours but if the students can finish it earlier, they can leave the class. The test is easy to administer and to assess (dominated multiple-choice and matching). The answer key is provided to assess the test.

Reliability

Brown (2004) points reliability is consistent and dependable. It means that if a teacher administers the same test with the same students in different time, a similar result should be guaranteed. Here, the oral test that is provided by National English Centre (NEC) is not reliable

because each student has a different ability to demonstrate the procedure and has different language abilities. Sometimes, the students are nervous about doing the oral test because they demonstrate in front of a general practitioner/ senior midwife. Therefore, they do not do the presentation well. On the other hand, they might feel fatigued in waiting their turn for an oral test. Moreover, no score list standard for the oral test makes the test is not reliability. The test has different criteria and interpretations if it is conducted by a different person.

The written test is reliable. The test will have the same result even if the facilitator gives it on different occasions because it is dominated with multiple-choice and matching. The answer key is supported to assess the test. The instructions and the written text in the problem sheet are clear. It can be understood and read clearly.

Validity

Validity is related to effectiveness and has an important principle. The written test that consists of listening, reading, language focus, and writing is not valid. Listening is given to assess students 'ability to listen. The topic is about first aid. It is one of the topics that the students learn during the teaching and learning. Reading is given to assess students 'ability in reading. The topic is about choking. It is one of the topics too. In addition, language focus is given to assess students 'ability in grammar, and writing test is given not to assess students' ability in writing. In fact, the language focus is low validity. It is because teaching and learning is only focused on learning about procedures, not in grammar. In addition, the writing only asks the students to match the procedure with the name of the procedure. The space for answering is tight so several students sometimes only write the alphabet with the complete answer. The content of the writing test is not valid. It is because in the writing part, most of the procedures are not included in the syllabus and material in teaching and learning, like Kegel exercise, tailor sitting, and making a fist. The students feel unfamiliar too with the procedures; they do not learn them during teaching and learning activities. On the other hand, the oral test is valid in terms of its content because match with the syllabus and the topic that are delivered during teaching and learning. It is about procedures, like how to bathe the baby, how to give intramuscular injections and how to do vulva hygiene.

Authenticity

Authenticity is related in real world context. It means that a test should be relevant in daily students' activities. Both oral test and written test of pre-intermediate level for midwife class are considered authentic since they represent real life activities as a midwife, like giving the first aid, how to bath the baby, baby chocking, and fever.

Washback

Washback is the effect of testing on teaching and learning (Brown, 2004). The effects of washback are positive and negative. The effect of oral test is positive because the students can get additional information from the general practitioner or senior midwife if the procedure that they did in the test is inappropriately. Thus, the students can learn their weakness from the test.

The written test has negative washback. The structures of the test are dominated by multiple choice and matching. It does not educate the students to think. They can easily to decide the correct answer of the test without thinking any longer. Five major points are concluded in a table below.

Table 2. Five major points of final test pre intermediate for midwife

	Oral Test	Written Test
Practicality	It is impractical: 1. It needs additional money to pay general practitioner/senior midwife.	It is practical: 1. It is cheaper. 2. The duration of the test is flexible.

	2. It needs extensive time. 3. It must take place in a medical laboratory. 4. It does not have specific and efficient evaluation procedure.	3. It easy to administer and assess because dominated with multiple choice and matching. 4. It has answer key so it is easy to score.
Reliability	It is not reliable because each student has different ability in presenting a procedure. Therefore the score and evaluation is different.	It is reliable because the result of the test will be same. It is dominated by multiple choices and matching. On the other hand, the instruction and the written text in the text are clear.
Validity	It is valid because the topics are related to the syllabus and teaching and learning activities.	It is moderate validity. The topic in listening and reading are related to the syllabus and teaching and learning activities while the topics in language focus and written test are not stated in both.
Authenticity	It is authentic because it is closely related to real life activities as midwife.	It is authentic because it is closely related to real life activities as midwife.
Washback	It has positive effect because the students can get additional information to improve their weakness in doing presentation.	It has negative effect because dominated with multiple choices and matching.

From the table above can be seen the strength and the weakness of final test for pre intermediate level. The strength is the test has assessed four skills that essential for language test, speaking, listening, reading, and writing. Furthermore, the test is in categorized of lower order thinking. It is as the weakness of the test. The five key terms in bloom taxonomy is remembering, understanding, applying, analyzing, and evaluating. The final test is stated in the second position which understands. The test only accommodates the students understanding by giving the multiple choices, matching, and asking the students to do presentation related the procedure that the students learn during teaching and learning activities while the level of the test is for the pre-intermediate level.

CONCLUSION

To sum up the final test of pre intermediate test for midwife does not fulfill all five criteria of a good test, namely practicality, reliability, validity, authenticity, and washback. The oral test contains three principles, which are validity, authenticity, and washback. The same goes to the written test since it covers validity, authenticity, and washback.

REFERENCES

- Ahsanu, M. (2013). Oral test: a powerful tool for assessing students' actual achievement in language learning. *Celt*, 13(1), 1–19.
- Brown, H. D. (2001). *Teaching by principle: An interactive approach to language pedagogy (second ed.)*. New York: Pearson Education.
- Brown, H. Douglas. (2004). *Language Assessment Principle and Classroom Practice*. Longman.
- Cresswell, J. . (1994). *Research Design: Qualitative and Quantitative Approaches*. SAGE Publications.
- Elthia, M. W., & Alawiyah, R. (2023). *An Analysis of the Quality Test Items Used in the Final Semester Assessment at MTs Nurul Yaqin Gili Air*. 06(01), 9638–9647.
- Fraenkel, J. R., & Wallen, N. E. (2009). *How to Design and Evaluate Research in Education (7th ed.)*. McGraw-Hil.
- Heaton, J. B. (1989). *Writing English Language Test: Longman Handbooks for Language Teachers*. Longman Inc.

- Latif, F. et al. (2013). *Critico-analytic Study of ESP Final Exam Tests for Students of Accounting in Iranian Universities*.
- Mahendra, E. (2023). Oral Assessment in Triggering Student Higher-Order Thinking Skills (Hots). *JISAE: Journal of Indonesian Student Assessment and Evaluation*, 9(1), 8–14.
<https://doi.org/10.21009/jisae.v9i1.32928>