

Promoting Assisted-Online Platform in Evaluating Reading Exercise in English for Academic Purpose Program

Khotimah Mahmudah

English Education Department
Teacher Training and Education Faculty
Universitas Sebelas Maret

Corresponding email: khotimah.mahmudah@yahoo.com

Ngadiso

English Education Department
Teacher Training and Education Faculty
Universitas Sebelas Maret

Nur Arifah Drajati

English Education Department
Teacher Training and Education Faculty
Universitas Sebelas Maret

Abstract:

Along with the development of the digital era, how to pass learning evaluation can be done by using one of platforms in technology. This is no exception in the case of evaluation for academic reading at the university level. This study used Google Form as a platform for conducting an evaluation process of reading exercise. The purposes were to create a chance for university students that join in English for academic purpose program in order to evaluate a specific material outside of class, to determine the levels of difficulty from reading exercises assisted-online platform, and to determine the discrimination indexes of reading exercises assisted-online platform. This paper was qualitative-content analysis research involved fifteen university students who came from different programs. The data collection was gathered from observation and document analysis. The finding reveals that the evaluation of reading exercise can be done effectively using online platform outside of class. This result is a novelty in doing evaluation for English Teacher. In conducting a test, by using an online platform, it also indicates that the test is valid and reliable. The implication is that the online platform has several benefits in helping teachers carrying out an evaluation or developing a test of a lesson and implementing the test outside of the class as effective as in the class.

Keywords: English for Academic Purposes; Evaluation; Reading Exercises; Online Platform

DOI : 10.20961/ijie.v2i2.26438



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

Introduction

In the digital era, many researchers have evaluated teaching-learning process which utilized a tool as the result of rapid technology development. The changing not only concerns how the teachers use the technology in evaluation but also select and match a tool to certain task with clarity and foresight (Levy, 2009). One of the results from many kinds of research is that the use of Interactive White Boards for the evaluation of reading activities in school, this research revealed that the reading test gave rise to a global indication (reading coefficient) on the efficiency of reading on digital devices (Diarra & Kubryk, 2014). The other research used 'Test/ef' Models into Evaluating a Reading Comprehension Test, and the result showed that these models could be used in the test development process to get the validation of the test (Ha, 2017). The evaluation was not only done for assessing students' achievement after joining the teaching-learning process but also assessing the teachers' performance during teaching and learning in the class. As the example, research was done involving several teachers as respondents, and the result showed that the teachers' performance in the stage of an opening lesson was indicated good enough in providing student learning motivation, and in the core learning activities or the elaboration section, teachers did not give the students a chance to express their ideas through meaningful activities and tasks. However, this previous studies did not analyze the evaluation test deeply in the context of ESP. In responding to this phenomena, the researchers in this study conducted the research related to the evaluation in the context of the ESP program.

Here, English for academic purpose (EAP) is a program in one of the universities in Surakarta for every student who wants to improve the ability in English, especially in reading English text for an academic situation. The function is to facilities students learning English in order to compete in academic context globally. However, a transitional stage between traditional teaching strategy without technology and modern teaching strategy with technology is quite involved in teaching EAP class. In order to make a changing, this paper aims to create a chance for non-language learner in university to evaluate a specific material outside of English Academic Purpose Class, to determine the levels of difficulty from reading exercises assisted-online platform, and to determine the discrimination indexes of reading exercises assisted-online platform.

Literature Review

The definition of educational evaluation

Gardner in Hadi (2009) provides the definition of educational evaluation as (1) evaluation as a professional judgment or decision, (2) evaluation as a measurement, and (3) evaluation as an assessment of conformity between achievement or outcome and goal, (4) decision which are evaluation-oriented, and (5) the objectives confronted with the evaluation. In addition, European Network of Quality Agencies states that an evaluation is the core activity in order to assure quality in higher education, and it means that doing the evaluation in the learning process is likely doing assessing teaching and academic studies in a subject or department and it is related to degree program (ENQA, 2012).

The consideration in determining a format test in a reading exercise

The type of test questions must be decided on the basis of several facts such as the university subject concerned, the purposes of the examination, the length and reliability of the proposed examination, preferences of teachers and pupils, the time available for the examination, whether factual knowledge or thinking which is tested (Ruch, 1924).

Before determining the format test, a teacher needs to consider some practical steps to test construction, it can be done with understanding the purposes of a test (assessing clear or ambiguous objectives) such as to provide a record for assigning grades, to provide a learning experience for students, to motivate students to learn, to serve as a guide for further study, to assess how well students are achieving the stated goals of the lesson, and to provide the instructor with an opportunity to reinforce the stated objectives and highlight what is important for students to remember (Brown, 2013).

Reading in the context of English for Academic Purposes

There are four types of reading for assessing and evaluation in teaching and learning; 1) Perceptive Reading, 2) Selective Reading, 3) interactive reading, 4) Extensive reading, and among of the types show the relationships of length, focus, and process (Brown, 2013).

For this study, the researchers used the design of interactive reading that is suitable for ESP context. In interactive reading, it focuses on parts of language in several paragraphs designed to be one page or more, and typical genres in interactive reading are anecdotes, short narratives and, descriptions, memos, announcements, recipes, etc. Also, an interactive task identifies relevant features such as lexical, symbolic, grammatical, and discourse within texts of moderately short length with the objective of retaining the information that is processed. It is also more focus on top-down processing than bottom-up. It depends on students' prior knowledge about English as a foreign language. Then, several designing assessment tasks for interactive reading are Cloze Tasks, Impromptu reading plus comprehension question, Short-answer tasks, Editing (Longer texts), ordering task and information transfer for reading charts, maps, graphs, Diagrams (Brown, 2013).

Method

This research used qualitative-content analysis method. This method is focused on studying and analyzing a communication in a systematic, objective, and quantitative manner for the purpose of measuring variables (Prasad, 2008, p. 2). Krippendorff (1980) in Prasad (2008) says that this method is also a research technique that focuses on making replicable and valid inference from data to their context. Based on the references, the researcher decided the objective of qualitative-content analysis in this research. The objective was to get information related to a current status of phenomena which used a technology in doing evaluation for EAP class. The researcher adopted six steps in doing qualitative content analysis: 1) Formulating the research questions and objectives, 2) Selecting content and sample, 3) Developing content categories, 4) Finalizing units of analysis, 5) preparing a coding schedule, pilot testing, and checking inter coder reliabilities, 6) analyzing the collected data (Stempel, 1989).

Participant

The study involved non-language learners that have taken the EAP exam as a university entrance requirement or have attended EAP class as one of the requirements for thesis examination. After the researcher got the respondents' information from the observation data, the researchers selected and asked the willing respondents to join in this research. Here, this research involved 15 respondents consisting of 8 females and seven males.

Data Collection and Analysis

This research used a qualitative technique to analyze the data. The qualitative technique was used to describe the data from ITEMAN analysis. ITEMAN is an instrument analysis program that requires data files in ASCII (text-only) format. All analyzed data should be inputted in a single file, and ITEMAN can analyze up to 750 items at a time, with the number of test takers virtually unlimited (Hadi S. , 2008).

a. Instrument

The researchers used an instrument to gather data from the respondents. The instrument was a set of reading text followed by some questions based on the specific indicators from English Academic Purpose Class Syllabus, and this instrument were filled by students from non-English program study who joined in the EAP class.

b. Procedure

The following is several steps to conduct this research in the field:

- ❖ First, the researchers did an observation about the target respondents, in which the researchers collected the respondent data about the experience as a test-taker for EAP test.
- ❖ Second, designing an instrument. The instrument consists of 24 questions related to academic reading in the context of English Academic Purpose for university students. Such as: providing descriptive text and recount text on the test. (*see figure 1 and 2*)
- ❖ Third, the researchers carried out testing or trying-out the questions from a group of university students, after getting the result, then editing the instrument.
- ❖ The fourth, the researchers distributed the instrument to the 15 respondents as an experiment for reading exercise assisted-Google Form.

- ❖ The fifth, the researchers analyzed instrument that was filled by students using ITEMAN Program.
- ❖ The sixth, the data of reading exercise were analyzed based on the problems and theories that were stated in this paper.

Text Type 1: Descriptive Text

Read the following text to answer questions number 1 to 4.

My Gorgeous Buddy

I have a special Buddy. She is my classmate and always sits next to me. Her name is Alysa. She is seventeen years old. She is extremely beautiful girl and very simple on the look. However, I really adore her. She is not only kind but also friendly.

Alysa comes from a poor family. Her father has passed away and her mother is a laborer in a mill. She has two young brother, but she does not have any sister. Her first young brother' name is Boy who is younger a year than Alysa. He works as waiter in a famous eatery. Her second young brother is Tommy who sells magazines after school, although He is just fourteen years old. While Alysa works part-time in school canteen.

One thing that I always admire about her is that she can manage her time well, and she always looks cheerful.

1. What is the purpose of the text? *

- a. to tell the writer life
- b. to describe the writer's special friend

Figure 1. The Illustration of Descriptive Text on Google Form

Text Type 2: Recount Text

Fill in the Blanks number (5) to (9) in the sentences below based the word(s) on the question number 5 to 9.

Budi who come from Kalimantan Island lived in Surakarta from 2014 to 2016. He (5) _____ around Universitas Sebelas Maret area. At the first time, He (6) _____ speak any Javanese language. He also (7) _____ any kinds of Javanese food, and he felt so hard to adapt with those kinds of food. It was (8) _____ that He was always homesick, but he just thought that he (9) _____ to survive finishing his study. |

Question Number 5 *

- a. is living
- b. was living
- c. live
- d. lived

Figure 2. The Illustration of Recount Text on Google Form

The Validity and Reliability

❖ Rule of Thumb for Results

In the ITEMAN program, there is the score of Alpha. It is The Reliability of Test Score. The following is A rule of thumb for interpreting alpha for dichotomous questions (i.e., questions with two possible answers) or Likert scale questions is (Stephanie, 2014):

Cronbach's alpha	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Findings and Discussions

Observation data: Students that did Reading Exercise assisted-Google Form

Table 1. The Data from Students that have EAP Exam Experiences

Study Program	University Students	Male	Female	Taking EAP class	Having an experience of EAP Exam
Indonesian	1	0	1	1	1
Science Education	5	2	3	3	5
Economic Education	5	2	3	4	5
Translation Linguistics	1	1	0	0	1
Education of History	1	1	0	0	1
Education of Technology	1	0	1	0	1
Mathematics education	1	1	0	0	1
	15	7	8	8	15

The levels of difficulty from reading exercises assisted-online platform

After analyzing Level of difficulty using ITEMAN, the equivalents below can be used in interpreting the level of difficulty from the data (Hadi S. , 2008):

0.00 – 0.29 = Very Difficult (or Difficult)

0.30 - 0.70 = Moderately Difficult (or Medium)

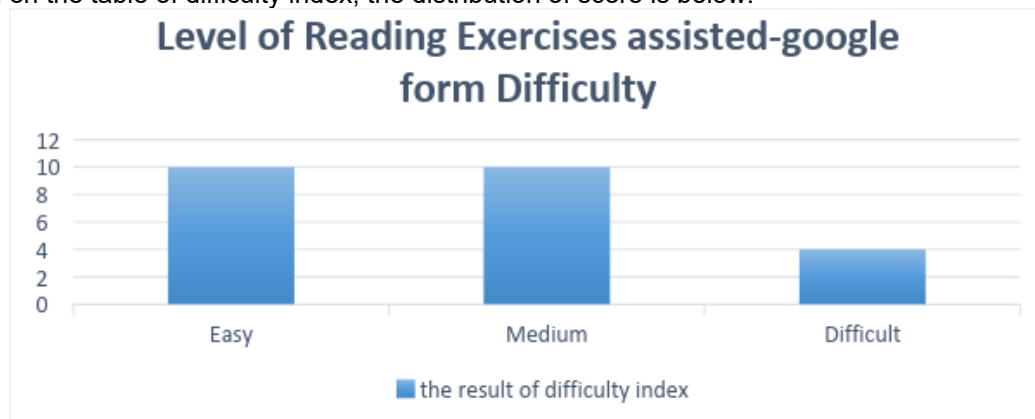
0.71 - 1.00 = Very Easy (or Easy)

Table 2. Analyzing the level of difficulty

Item	Level of Difficulty Score	Level of Difficulty
Item 1	0.933	Easy
Item 2	0.800	Easy
Item 3	0.867	Easy
Item 4	0.667	Medium
Item 5	0.200	Difficult

Item 6	0.467	Medium
Item 7	0.267	Difficult
Item 8	0.200	Difficult
Item 9	0.333	Medium
Item 10	0.467	Medium
Item 11	0.800	Easy
Item 12	0.600	Medium
Item 13	0.733	Easy
Item 14	0.933	Easy
Item 15	0.600	Medium
Item 16	0.667	Medium
Item 17	0.800	Easy
Item 18	0.733	Easy
Item 19	0.800	Easy
Item 20	0.667	Medium
Item 21	0.800	Easy
Item 22	0.600	Medium
Item 23	0.267	Difficult
Item 24	0.600	Medium

Based on the table of difficulty index, the distribution of score is below:



The result of difficulty index was healthy enough because most of the items were categorized as an easy and medium item. It means that most of students from different major could do this reading test in EAP class. Then, this reading test could be used for evaluating the students' reading skill in different condition of class.

The discrimination indexes of reading exercises assisted-Google Form

After analyzing discrimination index using ITEMAN, this data can be used in interpreting the discrimination index score as below (Hadi S. , 2008):

- 0.40 – up : Outstanding items and to be retained
- 0.30 – 0.39 : Good, but the items possibly subject to revision (minor revision)
- 0.20 – 0.29 : Marginal items, but the items usually need to revision (major revision)
- ≥ 0.19 : Poor items, and to be rejected

Table 3. Discrimination Index Result

<i>Item</i>	<i>The Discrimination Index Score</i>	<i>The Conclusion (retained/revised/rejected)</i>
<i>Item 1</i>	0.373	Minor Revision
<i>Item 2</i>	0.698	Retained
<i>Item 3</i>	0.145	Rejected
<i>Item 4</i>	0.464	Retained
<i>Item 5</i>	0.397	Minor Revision
<i>Item 6</i>	0.313	Minor Revision
<i>Item 7</i>	-0.037	Revised/discarded
<i>Item 8</i>	0.226	Major Revision
<i>Item 9</i>	0.580	Retained
<i>Item 10</i>	0.669	Retained
<i>Item 11</i>	0.048	Rejected
<i>Item 12</i>	0.413	Retained
<i>Item 13</i>	0.346	Minor Revision
<i>Item 14</i>	0.373	Minor Revision
<i>Item 15</i>	0.888	Retained
<i>Item 16</i>	0.406	Retained
<i>Item 17</i>	0.698	Retained
<i>Item 18</i>	0.841	Retained
<i>Item 19</i>	0.698	Retained
<i>Item 20</i>	0.377	Minor Revision
<i>Item 21</i>	0.698	Retained
<i>Item 22</i>	0.664	Retained
<i>Item 23</i>	0.581	Retained
<i>Item 24</i>	0.357	Minor Revision

Here, all items were questions that used to evaluate students' ability in reading. However, the evaluation was done using a Google Form as one of technological platform. Then, after testing the question, the

researcher calculated the discrimination index in order to get information which questions needed to be improved or deleted based on theory above. Based on table 3, it can be concluded that:

- There are 6 questions which should get Minor Revision
- There is just 1 question which should get Major Revision
- 14 questions should be retained or can be used for the other tests
- There are 2 questions that should be rejected
- The last, there is 1 question with minus DI score that should be totally revised or discarded

Test Reliability

A test cannot be valid unless it is reliable (Alderson, Clapham, & Dianne, 1995, p. 128). In this case, if the test does not measure something consistently, it follows that the test cannot always be measuring accurately what something measured. Besides, a test must have the predictive validity the developer would expect. Then, Henning, 1987 in (Alderson, Clapham, & Dianne, 1995) defines that validity as general refers to the appropriateness of a given test or any of test's parts as a measure of what it is recognized to measure, and the term "valid" is used to describe some purposes should usually accompany a test.

Based on the reference above, Alpha is the Reliability of Test Score. The alpha coefficient for the 24 items is (0.846), it means that the items have high internal consistency. A reliability coefficient of .70 or higher than α is considered "acceptable" in most social science research situations. It means that the reading test assisted-Google Form measures accurately what should be measured based on the learning objective.

Conclusion

This paper presented the factual data that gave the readers information about a platform that can be utilized in evaluating the teaching-learning process. The difficulty index is also indicated normal enough because most of the items are categorized as a natural and medium item. Based on the discrimination index result, six questions belong to Minor Revision category, just 1 question belong to Major Revision category, 14 questions were indicated to be retained, 2 questions should be rejected and the last 1 question with minus DI score that should be totally revised or discarded. Analyzing the discrimination index can make a teacher knowing the quality of the test items, which items need to be retained, revised, or rejected. So the teacher can evaluate the questions of the test after calculating the result of students' answer. The test was valid and reliable which means the test can be used for any other test for the same purposes. The implication is that Google Form has several benefits in helping teachers carrying out an evaluation or developing a test of a lesson and implementing the test outside of the class as effective as in the class.

Acknowledgments

The researcher would like to thank all participants who have contributed to implementing this research until it can be completed properly. The researcher also gives special honor thanks to Endowment Fund for Education (LPDP RI) who is being a sponsor giving a chance to the researcher to carry out the research.

Reference

- Alderson, J. C., Clapham, C., & Dianne, W. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Brown, H. D. (2013). *Language Assessment Principle and Classroom Practice*. San Francisco, California: Longman.
- Diarra, M., & Kubryk, G. (2014). The Use of Interactive White Boards for the Evaluation of Reading Activities in School. *International Conference: the Future of Education*.
- ENQA, E. N. (2012). *The Bologna Process, Glossary*. <http://www.bologna-berlin2003.de/en/glossary/>.
- Ha, D. T. (2017). The Implementation of Testlet Models into Evaluating a Reading Comprehension Test. *International Journal of Scientific & Engineering Research*.

- Hadi, S. (2008). *Analisis Tes/Instrumen dengan ITEMAN*. Yogyakarta: Lab. Komputer Universitas Negeri Yogyakarta.
- Hadi, S. (2009). *Faktor yang Mempengaruhi Kualitas Pembelajaran*. Yogyakarta.
- Levy, M. (2009). Technologies in Use for Second Language Learning. *The Modern Language Journal*, 769-782.
- M., V., Turunen, H., & Bondas, T. (2013). *Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study*. Wiley Publishing Asia Pty Ltd.
- Prasad, B. D. (2008). *Research Method for Social Work*. New Delhi: Lal Das, D.K and Bhaskaran.
- Ruch, G. (1924). *The Improvement of the Written Examination*. England: Oxford.
- Stempel, G. (1989). *Research Methods in Mass Communication*. Englewood Cliffs: Prentice-Hall.
- Stephanie. (2014). Statistic How To. <https://www.statisticshowto.datasciencecentral.com/cronbachs-alpha-spss>.

