Developing web-based reading tests for the students of English language education

Ulul Azmi

Universitas Nadhatul Ulama Yogyakarta. Jl. Lowanu No.47, Yogyakarta, 55162 Indonesia
* Corresponding author: E-mail: el_ziemmy@yahoo.com

Abstract

The objectives of this research are (1) to develop web-based reading tests for the second-semester students of English Language Education Department, Universitas Negeri Yogyakarta in the Academic Year of 2013/2014, and (2) to know the students' responses about the product that are developed in English Language Education Department, Universitas Negeri Yogyakarta in the Academic Year of 2013/2014. This is a research and development study which employed nine steps. The subjects of this study were the second-semester students of the English Language Education Department, Universitas Negeri Yogyakarta in the Academic Year of 2013/2014 that consisted of twenty students as small group evaluation and fifty-two students as field trial evaluation. The results of this research show that the validation result from the material expert shows that the average point is 3.95. It is categorized good. The validation result from the media expert shows that the average point is 4.00. It is categorized good. Therefore, it can be concluded that web-based reading tests are appropriate to be implemented in the Reading subjects. Furthermore, the result of the questionnaire of the students shows that the average point is 3.87. It is categorized good. It can be concluded that the students agree with developing web-based reading tests and there are no product revisions. The final product of this research referred to an online test system with domain address *www.wbrt-pbi-uny.com.*

Keywords: Test, web, reading

_____

INTRODUCTION

In the education, every learning activity needs to be measured in order to see the result of the learning achievement. It is called evaluation. The evaluation that is usually used for measuring is a test. Brown and Lee (2001, p. 16) states that test is assesement that could provide authenticity, motivation, and feedback to the students. The test can be designed with the result score provided for the students, thus it is able to know their ability and can improve it. In line with that, Wahyuni and Ibrahim (2012) mentions that the test is a tool to arrange the evaluation that is done by the students as test participant. The test usually produces achievement of the students in the teaching and learning process. In other words, the test is used to measure the skill, abilty, knowledge of the students. Nowdays, the model of the test still used paper-based test. In the paper-based test, the teacher and the students meet in the class and the teacher gives the students test. This test is good enough. However, the test need more fee to copy the test. By using technology, the teacher can use the modern test, such as using web-based test. She/he can avoid to more copying the test, the students will be accustomed to use the modern test. In order that, the students are not socially awkward in using the modern technology, such as using internet.

Internet can links computers all over the world into a single electronic communication network. It is usually called Web. Web is a system for acessing and viewing information on the internet. Brown (2004) states that web test has a specific type of computer-based test. It is a *Computer-Adaptive Test* (CAT). *Computer-Adaptive Test* is a set of questions that is received by examinee. Examinee gets the test according to her or his performance level. She or he starts to answer each question of the test and the computer scores the question. The computer is programmed to fulfill the test design that examinee sees only one question at a time, and the computer scores each question before selecting the next one. As a result, examinees cannot skip questions, and once they have entered and confirmed their

answers, they cannot return to questions or to any earlier part of the test. Besides, Li et al. (2013, p. 155) state that the internet use considered as a means of accessing resources for challenges and mass medium with the ability to fulfill personal communication and metiated needs. It can be concluded that the users can look for the information, make fun, look for a job, and look for what they need by using the internet.

In the teaching language, there are many models of the test that are expanded to measure language skills. The language skills are listening, speaking, writing, and reading. Those language skills are important in the teaching and learning process, especially in language learning. However, the reseacher will be discussed one of those language skills. It is reading. Reading is viewed as a tool of communication in written language through the form of the texts. It means that this skill is regarded as an efficient way to acquire information about many aspects of life. Djiwandono (1996, p. 62) mentions that reading is the important activity in this modern era. Many people can absorb a lot of information by reading. Many people enrich their knowledge and access information through reading books, magazines, and newspaper. Reading plays an important role and becomes a daily necessity for every person who wants to acquire whatever information he/ she needs. Reading does not take place without comprehension and comprehension cannot occur without thinking. Therefore, the reading process is closely related to thinking. Comprehension is the reason for reading. If the readers can read the words but do not understand what they are reading, they are not really reading. Reading is not only looking at words, phrases, and sentences, but also understanding the meaning. In line with that, Grabe (2009) states that reading can be easily defined simply as the ability to derive understanding from written text, the simple definition is denied by the complexity inherent in the reading ability. He believes that everyone has a combination of skills and abilities, they bring when they begin to read.

Reading subject in the English Language Education Department of Universitas Negeri Yogyakarta is given to the students in five semester (reading I until reading V). Reading subjects are different in every semester. It is given from basic level until advanced level. Celce-Murcia et al. (2020) states that teaching reading requires reading tasks to be done by the students to improve the reading skills in order to achieve their learning goals including academic goals and real-life goals. Hence, the teaching of reading covers the students' learning goals. In other words, it is not only for the academic goals, reading instruction also plays a role for giving students' experiences to use their skills in real-life activities. To know the competence of the students in reading subject, the lecture usually gives the students a test to measure their ability in reading. In this case, the reading test in the English Language Education Department of Universitas Negeri Yogyakarta still used paper-based reading test. The lecture gives the students test, then the lecture can look at the students in doing the test immediately. Actually, there is no problem in this test. However, this test is less effective and efficient because the lecture need more time to give scores of the test manually for every student, and need much money to copy the test that depends on the number of the students. Web based tests are not developed yet in the English Language Education Department of Universitas Negeri Yogyakarta and the lectures do not use the network yet in Universitas Negeri Yogyakarta optimally.

Based on the problems above, the reseacher will develop web-based reading tests for the second semester students in the English Language Education Department of Universitas Negeri Yogyakarta in the academic year of 2013/2014. It is because the technology develops rapidly. There are some benefits using the modern technology, the students can access and can view much information, the students can practice and can learn the test by using internet, and the students can be accustomed in using the modern technology. The objectives of this research are to develop web-based reading tests for the second semester students of English Language Education Department, Universitas Negeri Yogyakarta in the academic year 2013/2014 and to know the students' responses about the product that is developed in English Language Education Department, Universitas Negeri Yogyakarta in the academic year 2013/2014.

Developing web-based reading tests are new issue in the English Language Education Department of Universitas Negeri Yogyakarta for the second semester students. There are some relevant studies that conducted in various developed English skills. One of the research studies is conducted by Luthfa (2013). Luthfa (2013) conducted a study on developing web-based reading

materials for eighth grade students of SMP 2 Playen Gunungkidul. She conducted her study in SMPN 2 Playen Gunungkidul. The aim of her study was to develop a set of web-based reading materials for eighth grade students of SMP 2 Playen Gunungkidul. The result of her research study was a proper set of web-based reading materials for the eighth-grade students of SMPN 2 Playen Gunungkidul. The next was Faqih's study (Faqih, 2013). He conducted a study on developing a website as a writing instruction for the tenth-grade students of Senior High School in the first semester. He conducted his study in Senior High School. The aim of his research was to develop a website as a writing instructionfor the tenth grade students of Senior High School in the first semester.The results of the developed website are called Moodle-Based Writing Website for the tenth grade students of senior high school in the first semester of which the address was at URL www.novian-faqih.com/moodle. Furthermore, Sariwulan (2013) concucted a study on developing computer-assisted language learning (CALL)-based materials for reading comprehension for Grade VIII students of SMP N 1 Prembun, Kebumen. The subjects of the research were students of Grade VIII of SMP N 1 Prembun, Kebumen Class VIII A and Class VIII B. Questionnaires and the interview guide line were the instruments used in the research to obtain the data. The data from the questionnaires were analyzed quantitatively and the data from the interview were investigated qualitatively. The result of the needs analysis showed that the students needed materials for reading comprehension which were attractive and interesting and based on their study background.

Therefore, the researcher develops web-based reading tests for the second semester students of English Language Education Department, Universitas Negeri Yogyakarta in the academic year of 2013/2014 in this research. There are some similarities and differences between those studies above and this study. The similarities are to develop the products, to develop especially in reading skill, and to develop web-based. The differences of those study above are to develop products of the materials and the products are developed for the junior high school students or senior high school students. Whereas, this study is to develop the product of the tests and the product is developed for the second semester students of English Language Education Department, Universitas Negeri Yogyakarta in the academic year of 2013/2014.

METHOD

The main purpose of this study is to develop web-based reading tests for the second semester students of English Language Education Department, Universitas Negeri Yogyakarta in the Academic Year of 2013/2014, which can be classified as a product. The type of this study is Research and Development (R&D). Research and Development (R&D) is a development model in which findings are used to develop new products (Gall et al., 2003, p. 569).

The subjects of the try-outs are the second semester students of English Education Department, Universitas Negeri Yogyakarta taking Reading 2 course. They consist of twenty students and fifty-two students of the second semester of English Education Department, Universitas Negeri Yogyakarta.

This research follows the major steps of R & D cycle as suggested by Gall et al. (2003, p. 570) which are as follows: (1) research and information collecting, (2) planning, (3) develop preliminary field form of product, (4) preliminary field testing, (5) main product revison, (6) main field testing, (7) operational product revision, (8) operational field testing, (9) final product revision, (10) dissemination and implementation (Diani et al., 2018; Dwianto et al., 2017).

In line with that, Sugiyono (2016, p. 409) mentioned the steps of R & D, those are as follows: (1) identification of the problem, (2) information collecting, (3) product design, (4) design validation, (5) design revision, (6) field testing of the product, (7) product revision, (8) operational field testing, (9) final product revision, (10) design the final product.

From the steps of R & D stated by the experts above, It is found that there are some similarities about the steps of R & D which have been stated by those experts. Based on the similarities, the researcher combines with adaptation those steps into nine. Those are as follows: (1) planning and designing, (2) developing the first product, (3) validating the first product, (4) revising the first product,

(5) try-out the first product and revision, (6) developing the final product, (7) validating the final product, (8) revising the final product, (9) try-out the final product and revision.

The several steps above are able to be explained in the following below:

Planning and designing

In this step, the researcher analyzes the existing syllabus of Reading 2 in the English Language Education Department, Universitas Negeri Yogyakarta and creates the blueprint of reading test that is produced based on the analyis of the syllabus of Reading 2. Then, the blueprint is developed become the reading test.

Developing the first product

In this step, the researcher develops the reading test as the first product. This reading test is developed from the blueprint.

Validating the first product

In this step, the reading test as the first product is validated by the material expert. The suggestions and comments of the material expert are became the input to the next step.

Revising the first product

In this step, the researcher revises the reading test that is validated by the material expert. The suggestions and comments of the material expert are became the input for revising the reading test.

Try-out the first product and revision

In this step, the researcher conducts the try-out to know the validity and the reliability of the test. The researcher conducts the try-out for the twenty of the second semester students of English Language Education Department, Universitas Negeri Yogyakarta as small group evaluation and analyzes the result of the try-out. The researcher analyzes the validity of the test using *Point-Biserial Correlation* formula and the reliability of the test using *Kuder Richardson-20* formula. Then, the result of this step is became into the basis for doing the next step.

Developing the final product

Before developing web-based reading tests as the final product, the researcher creates the flowchart and the storyboard first to know the first design of web-based reading tests. Then, the researcher develops web-based reading tests.

Validating the final product

In this step, web-based reading tests as the final product are validated by the media expert. The suggestions and comments of the media expert are became the input to the next step.

Revising the final product

In this step, the researcher revises web-based reading tests that are validated by the media expert. The suggestions and comments of the media expert are became the input for revising the web-based reading tests.

Try-out the final product and revision

In this step, the researcher conducts the final try-out. The researcher conducts the try-out for fifty-two of the second semester students of English Language Education Department, Universitas Negeri Yogyakarta as field trial evaluation. In the end of this step, the researcher gives questionnaire to the students. The result of the questionnaire is became the input to revise web-based reading tests and to know the students' responses.

Techniques and Instruments of Data Collection

The techniques used in this research are the reading test and questionnaire.

The Reading Test

The test is an objective test that consist of 50 items with 4 choices. The scores techniques in the reading test are 1 and 0. Each correct answer is scored 1 and each wrong answer is scored 0. The possibility of the highest score obtained by students in the reading test is 50 and the possibility of the lowest score is 0. The researcher adapts the texts of reading test from Cambridge. The researcher creates the questions of the reading test as twenty seven (1, 2, 4, 5, 6, 7, 9, 10, 12, 19, 20, 21, 23, 25, 26, 27, 28, 30, 31, 32, 33, 34, 35, 41, 43, 45, 50), and adapts the questions of the reading test as twenty three (3, 8, 11, 13, 14, 15, 16, 17, 18, 22, 24, 29, 36, 37, 38, 39, 40, 42, 44, 46, 47, 48, 49).

Table 1. Categorization of the Reading Test

Category	Score
High	1
Low	0

Questionnaire

The questionnaire is given to the students to know the quality of the web performance. The questionnaire is used to support the result of the test. The questionnaire consisted of 20 items. The scoring technique of those options is 5-4-3-2-1. Thus, the lowest score of the item is 1 and the highest score is 5. Each item consists of a statement, followed by five options: 5. SS (Strongly Agree); 4. S (Agree); 3 AS (Rather Agree); 2. TS (Disagree); 1. STS (Strongly Disagree).

Table 2. Categorization of the Questionnaire

Category	Score
Strongly Agree (SS)	5
Agree (S)	4
Rather Agree (AS)	3
Disagree (TS)	2
Strongly Disagree (STS)	1

Techniques of Data analysis

Brown (2004, p. 387) defines validity as the degree to which the data collection procedure measures what it intends to measure. To analyze the data from reading test, the reseacher used validity and reliability of the test. To establish the validity of the test, the reseacher uses the item validity. In the item validity, the researcher test the validity of each item of the reading test using an item analysis called as *Point*-Biserial *Correlation*. The function of *Point-Biserial Correlation* is to determine the level of students' ability in doing a test, especially doing reading test and to determine the validity of test itself. Besides, employing the validity of the data, the researcher also employed item difficulty/facility and item discrimination. Then, the test should be measure the reliability of the test. The test should be measure the reliability of the test. The test itself should yield similar result (Brown, 2004). The researcher employs *Kuder Richardson-20* formulation to measure the reliability of reading test.

Besides, to analyze the data from the questionnaire, the reseacher used descriptive statistics. Descriptive statistics according to (Brown & Lee, 2001, p. 144) refers to a set of procedures that are used to describe or characterize the answer of a group of respondents to numerically coded questions. It will be converted to the qualitatative data. Meanwhile, it is classified into five categories. The categories are presented in Table 3.

Table 3. Conversion Table

Score	Interval	Category
5	$X > Xi + 1.80\ Sbi$	Very Good
4	$Xi + 0.60\ SBi < X \leq Xi + 1.80\ SBi$	Good
3	$Xi - 0.60\ SBi < X \leq Xi + 1.80\ SB$	Fair

| 2 | $Xi - 0.60\ SBi < X\ \leq Xi - 1.80\ SB$ | Poor |
| 1 | $X \leq Xi - 1.80\ Sbi$ | Very Poor |

Based on the conversion formula (Table 3), the qualitative data should be changed into quantitative data. The conversion can be seen in the Table 4.

Table 4. Conversion of Quantitative Data to Qualitative Data

Score	Interval	Cateogry
5	$X > 4.21$	Very Good
4	$3.40 < X \leq 4.21$	Good
3	$2.60 < X \leq 3.40$	Fair
2	$1.79 < X \leq 2.60$	Poor
1	$X \leq 1.79$	Very Poor

The conversion Table 4 is used to determine the proper or suitable category or not of the product that is developed. Developing product of web-based reading tests are called proper or suitable, if the result of the product has minimal criteria of Good or B category. The proper or suitable product is able to use as media or learning source.

RESULTS AND DISCUSSION

Before the product was tried-out, the product was validated by some experts. Those experts involved in giving comments and suggestions of the instrument evaluations. Those expert were as follows.

Material Validation

Data of Material Validation

Before the material was tried-out, it was validated by the expert. The material expert was A. Ghani Johan, M.Ed. The expert was validated the content, grammar, appearance of reading test. There were 50 questions of reading test that was validated by the expert. The researcher gave instrument evaluation to the expert. In order that, the expert gave some comments and suggestions for it. This was done on 10th of April 2014 in the Language Center.

There were some comments or suggestions to be revised by the reseacher. They were some questions should be revised before the researcher conducted the try-out. And also the reseacher gave some indicators of material aspect to the expert. In order that, the expert could give assessment scale for the indicators.

Result of Material Validation

From the result of the assessment, the score of the assessment is 3.95. It can be concluded that the reading test was developed by the reseacher is good because the assessment in the frequecy of 3.40 < X ≤ 4.19. It is presented in Table 5 and also it can be seen in the Figure 1.

Try-out of the first product

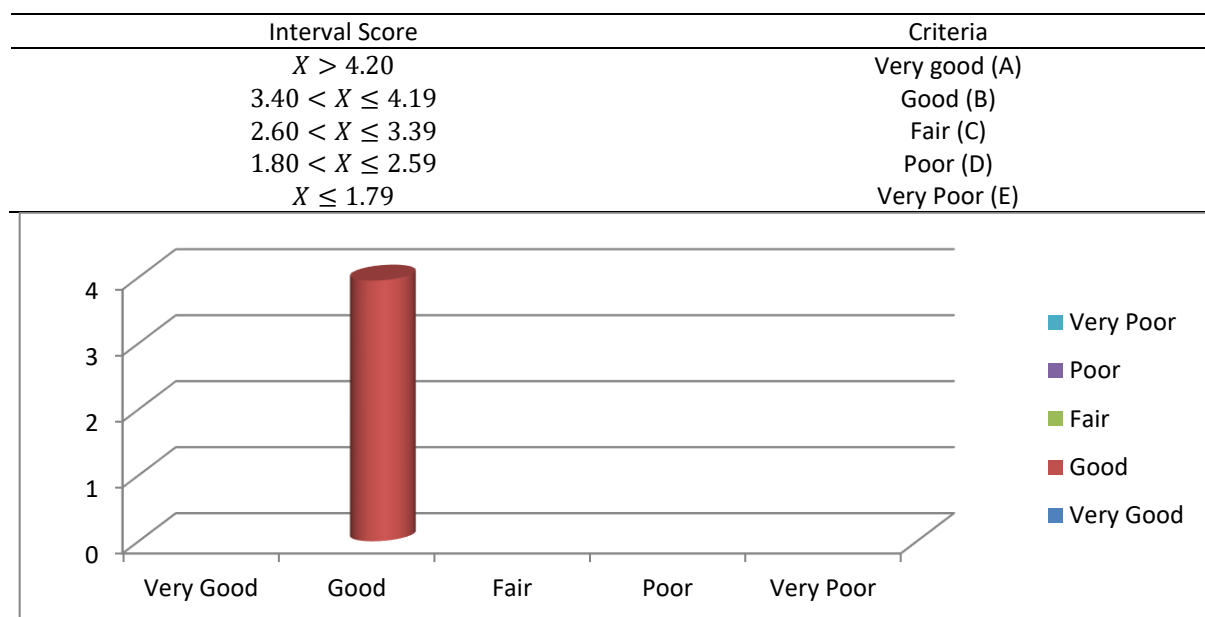Data of try-out the first product

In the try-out, the researcher gave 50 item questions of reading test and 60 minutes with the different type of the questions but still in the similar topics. The researcher developed the reading tests based on the syllabus of reading two in English Language Education Department of Universitas Negeri Yogyakarta. The researcher took 20 students of second semester in English Language Education Department of Universitas Negeri Yogyakarta for doing the reading tests. This was done on 14th of April 2014 in the English Language Education Department of Universitas Negeri Yogyakarta.

Result of item analysis

After the reading test was tried-out, it was found that there were 10 questions invalid. Those numbers were 5, 10, 14, 19, 22, 26, 32, 42, 43, and 47. The researcher used *Point-Biserial Correlation*

to know the validity of the reading test. Furthermore, the researcher used *Kuder-Richardson-20* to know the reliabilty of the reading test. The reliabilty of the reading test was 0.924.

Table 5. Result of Material Validation

Interval Score	Criteria
$X > 4.20$	Very good (A)
$3.40 < X \leq 4.19$	Good (B)
$2.60 < X \leq 3.39$	Fair (C)
$1.80 < X \leq 2.59$	Poor (D)
$X \leq 1.79$	Very Poor (E)

Figure 1. Diagram of Material Validation

Table 6. Validity of the Test

No.	Upper group	Lower group	Point Bisserial	Status
1.	6	4	0.58	Valid
2.	6	2	0.75	Valid
3.	5	3	0.42	Valid
4.	6	3	0.68	Valid
5.	5	4	0.20	Invalid
6.	6	4	0.58	Valid
7.	6	4	0.58	Valid
8.	5	2	0.49	Valid
9.	5	1	0.67	Valid
10.	6	6	0.00	Invalid
11.	5	3	0.42	Valid
12.	6	4	0.58	Valid
13.	5	2	0.49	Valid
14.	2	1	0.27	Invalid
15.	6	4	0.58	Valid
16.	5	3	0.42	Valid
17.	6	2	0.75	Valid
18.	6	3	0.68	Valid
19.	5	4	0.20	Invalid
20.	6	4	0.58	Valid
21.	4	2	0.33	Valid
22.	1	1	0.00	Invalid
23.	5	2	0.49	Valid
24.	3	1	0.42	Valid
25.	5	1	0.67	Valid
26.	2	2	0.00	Invalid
27.	5	3	0.42	Valid
28.	6	4	0.58	Valid
29.	5	1	0.67	Valid

No.	Upper group	Lower group	Point Bisserial	Status
30.	6	3	0.68	Valid
31.	6	2	0.75	Valid
32.	6	6	0.00	Invalid
33.	4	2	0.33	Valid
34.	6	3	0.68	Valid
35.	6	3	0.68	Valid
36.	3	1	0.42	Valid
37.	5	1	0.67	Valid
38.	6	4	0.58	Valid
39.	6	4	0.58	Valid
40.	6	4	0.58	Valid
41.	6	4	0.58	Valid
42.	5	4	0.20	Invalid
43.	2	1	0.27	Invalid
44.	5	3	0.42	Valid
45.	4	1	0.34	Valid
46.	6	3	0.68	Valid
47.	6	6	0.00	Invalid
48.	6	4	0.58	Valid
49.	6	4	0.58	Valid
50.	6	3	0.68	Valid

Media Validation

Data of Media Validation

After the material was tried-out to the students, the reseacher developed media. It was web-based reading tests and it was validated by the expert before it was tried-out. The expert was Herman Dwi Surjono, Ph.D. The expert was validated the content, appearance of web-based. The researcher gave instrument evaluation to the expert. In order that, the expert gave some comments and suggestions for it. This was done on 13th of Mei 2014 in the Instructional Technology Study Program.

There were some comments, suggestions, and opinions to be revised by the reseacher. They were the identity of the researcher in the home page should be completed, the choices should be A, B, C, D (not 1,2,3,4), the appearance of home page should be revised, every appearances of function should be given subtitles. And also the reseacher gave some indicators of appearance and system aspect to the expert. In order that, the expert could give assessment scale for the indicators.

Result of Media Validation

From the result of the assessment, the score of the assessment is 4.00. It can be concluded that the web-based test was developed by the reseacher is good because the assessment in the frequecy of $3.40 < X \leq 4.19$. It is presented in Table 7 and also it can be seen in the Figure 2.

Try-out of the final product

Data of try-out of the final product

In the try-out, the researcher took 52 students of second semester of English Language Education Department of Universitas Negeri Yogyakarta for doing web-based reading tests. This was done on 20th – 28th of May 2014 in the laboratory of English Language Education Department of Universitas Negeri Yogyakarta. The reseacher gave the students 40 questions of reading tests using web-based and 50 minutes with different type of questions but still in the similar topics. Web-based reading tests were developed that the students could not return to the previous questions and the students could divide the time of each questions. In the end of the test, the students were doing questionnaire to know the responses of the students about the product.
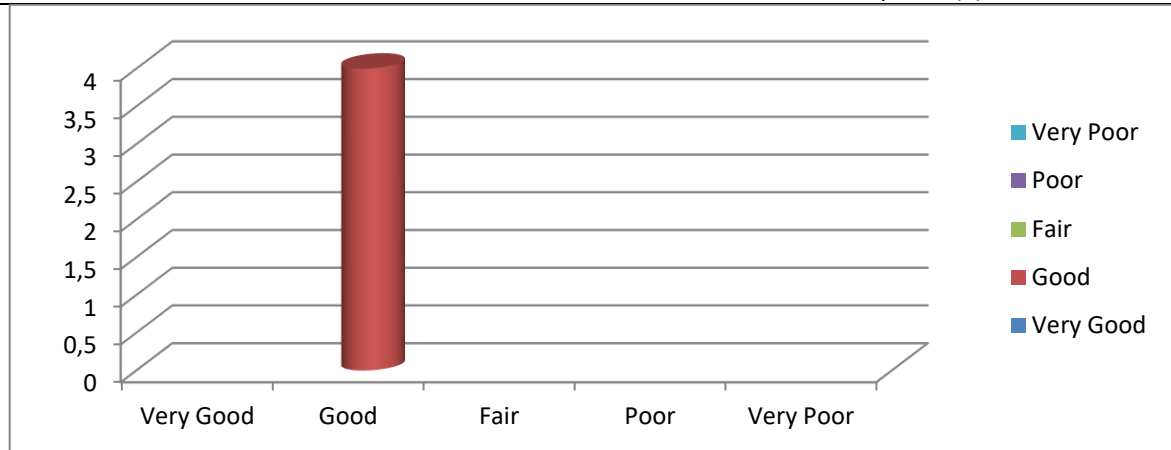
Results of the analysis

It was found that web-based tests were developed by the reseacher are good because the assessment of the students is 3.87. It is in the frequecy of 3.40 < X ≤4.19. It is presented in Table 8.

Table 7. Result of Media Validation

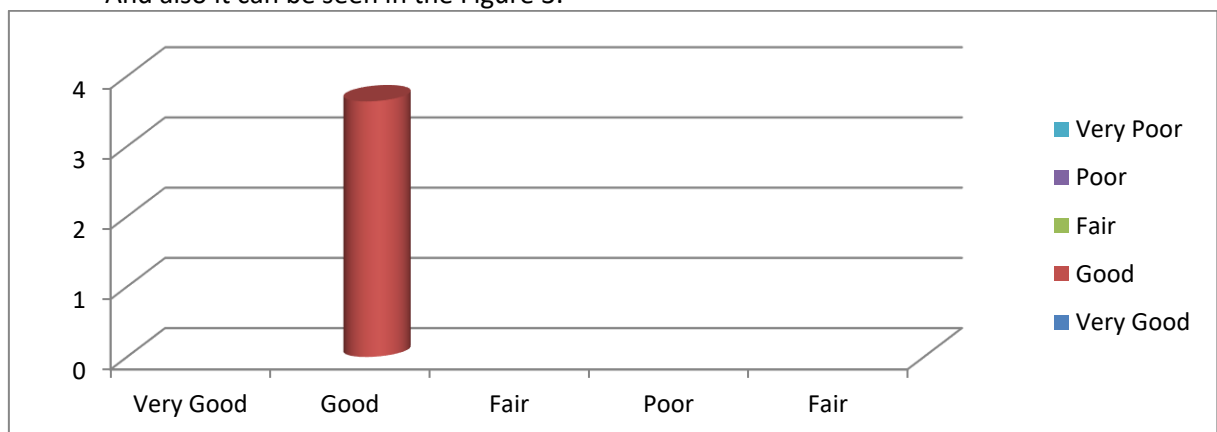| Interval Score	Criteria
$X > 4.20$	Very good (A)
$3.40 < X \leq 4.19$	Good (B)
$2.60 < X \leq 3.39$	Fair (C)
$1.80 < X \leq 2.59$	Poor (D)
$X \leq 1.79$	Very Poor (E)

Figure 2. Diagram of Media Validation

Table 8. Results of Questionnaire

Interval Score	Criteria
$X > 4.20$	Very good (A)
$3.40 < X \leq 4.19$	Good (B)
$2.60 < X \leq 3.39$	Fair (C)
$1.80 < X \leq 2.59$	Poor (D)
$X \leq 1.79$	Very Poor (E)

And also it can be seen in the Figure 3.

Figure 3. Diagram of Questionnaire

The assessment of the students can be described in the following below:

The program is easy to be used

In the indicator of "the program is easy to be used", there were 12 of the students (23.2%) answer "very good" criteria, 28 of the students (53.8%) answer "good" criteria, 10 of the students

(19.2%) answer "fair" criteria, and 2 of the students (3.8%) answer "poor criteria". Therefore, the average of this indicator is 3.96. It means that this indicator is "good" criteria.

The menu program is easy to be choosen

In the indicator of "menu program is easy to be choosen", there were 4of the students (7.7%) answer "very good" criteria, 32 of the students (61.5%) answer "good" criteria, 11 of the students (21.2%) answer "fair" criteria, and 5 of the students (9.6%) answer "poor criteria". Therefore, the average of this indicator is 3.67. It means that this indicator is "good" criteria.

The program has appropiate colour

In the indicator of "the program has appropiate colour", there were 5 of the students (9.6%) answer "very good" criteria, 30 of the students (57.7%) answer "good" criteria, 11 of the students (21.2%) answer "fair" criteria, and 6 of the students (11.5%) answer "poor criteria". Therefore, the average of this indicator is3.65. It means that this indicator is "good" criteria.

Register is easy to be used

In the indicator of "register is easy tobe used", there were 29 of the students (55.8%) answer "very good" criteria, 21 of the students (40.4%) answer "good" criteria, and 2 of the students (3.8%) answer "fair" criteria. Therefore, the average of this indicator is 4.52. It means that this indicator is "very good" criteria.

Login is easy to be used

In the indicator of "login is easy to be used", there were 27 of the students (51.9%) answer "very good" criteria, 22 of the students (42.3%) answer "good" criteria, and 3 of the students (5.8%) answer "fair" criteria. Therefore, the average of this indicator is 4.48. It means that this indicator is "very good" criteria.

The program has appropriate font

In the indicator of "the program has appropriate font", there were 1of the student (1.9%) answer "very good" criteria, 32 of the students (61.5%) answer "good" criteria, 16 of the students (30.8%) answer "fair" criteria, and 3 of the students (5.8%) answer "poor criteria". Therefore, the average of this indicator is 3.58. It means that this indicator is "good" criteria.

The appropriate size of the text

In the indicator of "the appropriate size of the text", there were 1 of the student (1.9%) answer "very good" criteria, 27 of the students (51.9%) answer "good" criteria, 17 of the students (32.6%) answer "fair" criteria, and 7 of the students (13.6%) answer "poor criteria". Therefore, the average of this indicator is 3.4. It means that this indicator is "good" criteria.

The interesting home page design

In the indicator of "the interesting home page design", there were 1 of the student (1.9%) answer "very good" criteria, 35 of the students (67.3%) answer "good" criteria, 13 of the students (25%) answer "fair" criteria, and 3 of the students (5.8%) answer "poor criteria". Therefore, the average of this indicator is3.63. It means that this indicator is "good" criteria.

The instruction of the test is clear

In the indicator of "instruction of the test is clear", there were 2 of the students (3.9%) answer "very good" criteria, 41 of the students (78.8%) answer "good" criteria, 8 of the students (15.4%) answer "fair" criteria, and 1 of the student (1.9%) answer "poor criteria". Therefore, the average of this indicator is 3.87. It means that this indicator is "good" criteria.

The appropriate time of the test

In the indicator of "appropriate time of the test", there were 6 of the students (11.5%) answer "very good" criteria, 18 of the students (34.6%) answer "good" criteria, 13 of the students (25%) answer "fair" criteria, and 15 of the students (28.9%) answer "poor criteria". Therefore, the average of this indicator is 3.29. It means that this indicator is "fair" criteria.

The program has good quality test

In the indicator of "program has good quality test", there were 4 of the students (7.8%) answer "very good" criteria, 41 of the students (78.8%) answer "good" criteria, 6 of the students (11.5%) answer "fair" criteria, and 1 of the student (1.9%) answer "poor criteria". Therefore, the average of this indicator is 3.88. It means that this indicator is "good" criteria.

The program is easy to learn for dividing time and doing test

In the indicator of "the program is easy to learn for dividing time and doing test", there were 8 of the students (15.4%) answer "very good" criteria, 35 of the students (67.2%) answer "good" criteria, 3 of the students (5.8%) answer "fair" criteria, and 6 of the students (11.6%) answer "poor criteria". Therefore, the average of this indicator is 3.87. It means that this indicator is "good" criteria.

The program is given the score result of the test quickly

In the indicator of "the program is given the score result of the test quickly", there were 20 of the students (38.5%) answer "very good" criteria, 30 of the students (57.7%) answer "good" criteria, and 2 of the students (5.8%) answer "fair" criteria. Therefore, the average of this indicator is 4.31. It means that this indicator is "very good" criteria.

There is no web-based reading test in our department

In the indicator of "there is no web-based reading test in our department", there were 18 of the students (34.6%) answer "very good" criteria, and 34 of the students (65.4%) answer "good" criteria. Therefore, the average of this indicator is 4.35. It means that this indicator is "very good" criteria.

The test that given is appropiate for the second semester students

In the indicator of "the test that given is appropiate for the second semester students", there were 12 of the students (23.1%) answer "very good" criteria, 28 of the students (53.9%) answer "good" criteria, 8 of the students (15.3%) answer "fair" criteria, and 4 of the students (7.7%) answer "poor criteria". Therefore, the average of this indicator is 3.9. It means that this indicator is "good" criteria.

The test is too easy to learn

In the indicator of "the test given is too easy to learn", there were 13 of the students (25%) answer "very good" criteria, 24 of the students (46.2%) answer "good" criteria, 4 of the students (7.7%) answer "fair" criteria, and 11 of the students (21.1%) answer "poor criteria". Therefore, the average of this indicator is 3.69. It means that this indicator is "good" criteria.

Time for the test is too limited

In the indicator of "time for the test is too limited", there were 2 of the students (3.8%) answer "very good" criteria, 33 of the students (63.5%) answer "good" criteria, 6 of the students (11.6%) answer "fair" criteria, and 11 of the students (21.1%) answer "poor criteria". Therefore, the average of this indicator is 3.48. It means that this indicator is "good" criteria.

This reading test is appropriate with the syllabus

In the indicator of "this reading test is appropriate with the syllabus", there were 13 of the students (25%) answer "very good" criteria, 36 of the students (69.2%) answer "good" criteria, and 3 of the students (5.8%) answer "fair" criteria. Therefore, the average of this indicator is 4.13. It means that this indicator is "good" criteria.

The text and the questions are clear enough

In the indicator of "the text and the questions are clear enough", there were 8 of the students (15.5%) answer "very good" criteria, 39 of the students (74.9%) answer "good" criteria,4 of the students (7.7%) answer "fair" criteria, and 1 of the student (1.9%) answer "poor" criteria. Therefore, the average of this indicator is 4.02. It means that this indicator is "good" criteria.

The questions of the test are easy enough

In the indicator of "the questions of the test are easy enough", there were 6 of the students (11.6%) answer "very good" criteria, 31 of the students (59.6%) answer "good" criteria, 4 of the students (7.7%) answer "fair" criteria, and 11 of the students (21.1%) answer "poor" criteria. Therefore, the average of this indicator is 3.63. It means that this indicator is "good" criteria.

CONCLUSIONS

The conclusions of the research are based on the results of web-based reading tests product. These are as follows: (1) Developing the product of the research was applied based on the combination of Borg and Gall concept and Sugiyono concept. There are nine stages to develop web-based reading tests product. They are planning and designing, developing the first product, validating the first product, revising the first product, try-out the first product and revision, developing the final product, validating the final product, revising the final product, try-out the final product and revision. It was found that the validation result from the material expert shows that the average point is 3.95. It is categorized good and the validation result from the media expert shows that the average point is 4.00. It is categorized good. Therefore, these can be concluded that web-based reading test are appropriate to be implemented in Reading subject. (2) Based on the students' responses, web-based reading tests as the product are good because the assessment score is 3.87. Therefore, web-based reading tests that are developed by the researcher does not need to be revised. It is proper to used.

REFERENCES

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Longman: Pearson Education, Inc.

Brown, H. D., & Lee, H. (2001). *Teaching by principles: An interactive approach to language pedagogy*. Prentice Hall Regents Englewood Cliffs, NJ.

Celce-Murcia, M., Brinton, D. M., & Snow, M. A. (2020). *Teaching English as a second or foreign language*. Cengage Learning.

Diani, R., Yuberti, Y., & Syarlisjiswan, M. R. (2018). Web-enhanced course based on problem-based learning (PBL): Development of interactive learning media for basic physics II. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *7*(1), 105. https://doi.org/10.24042/jipfalbiruni.v7i1.2849

Djiwandono, M. S. (1996). *Tes bahasa dalam pengajaran*. ITB Bandung.

Dwianto, A., Wilujeng, I., Prasetyo, Z. K., & Suryadarma, I. G. P. (2017). The development of science domain based learning media which is integrated with local potention to improve science process skill and scientific attitude. *Jurnal Pendidikan IPA Indonesia*, *6*(1). https://doi.org/10.15294/jpii.v6i1.7205

Faqih, N. (2013). *Developing a moodle-based writing website for the tenth grade students of the senior high school in the first semester*. Universitas Negeri Yogyakarta.

Gall, M. D., Gall, J. P., Borg, W. R. D., & Gall, J. P. (2003). *Educational research: An introduction* (7th ed.). Pearson Education Inc.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Ernst Klett Sprachen.

Li, J. Q., Liu, X., Wei, T., & Lan, W. (2013). Acculturation, internet use, and psychological well-being among Chinese international students. *Journal of International Students*, *3*(2), 155–166.

Luthfa, R. N. (2013). *Developing web-based reading materials for the eighth grade students of SMPN*

2 Playen Gunungkidul. Universitas Negeri Yogyakarta.

Sariwulan, F. (2013). *Developing computer-assisted language learning (call)-based materials for reading comprehension for grade VIII students of SMPN 1 Prembun, Kebumen* [Universitas Negeri Yogyakarta]. https://eprints.uny.ac.id/19233/

Sugiyono, S. (2016). *Metode penelitian pendidikan (Pendekatan kuantitatif, kualitatif, dan R & D)*. Alfabeta, CV.

Wahyuni, S., & Ibrahim, S. (2012). *Asesmen pembelajaran bahasa*. Refika Aditama.