

CHAPTER IV

FINDINGS AND DISCUSSION

In this chapter, the researcher presents the findings and discussion of the analysis related to the validity and reliability of summative English for ninth grade students at SMPN 3 Sungai Pandan.

A. Findings

1. Qualitative Phase

a. The Validity of the English Summative Test

SMPN 3 Sungai Pandan conducts a summative test every year for all subjects to determine students who pass the final level of junior high school. In English subject, the books that are used as references in the learning process are divided into 1 book title for students and 2 book titles for teachers, and also teachers usually use additional material from the internet, both from various learning portals such as Rumah Belajar, Ruang Guru, Youtube, etc. English teachers use learning methods during teaching such as using Project Based Learning, Discovery Learning, and Problem Based Learning.

The materials tested are Greeting Card, Advertisement, Label, Procedure Text, and Report Text. The type of material tested is described as follows:

1. Item number 1 is a greeting card text type with a purpose of the text question type.
2. Item number 2 is a greeting card text type with general inference question types. Inference questions are questions that are required to use clues in the text

to conclude what the text is talking about, in other words the questions are implicit.

3. Item number 3 is a text advertisement type with the type of inference question but more specifically the target is the audience.
4. Item number 4 is a text advertisement type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
5. Item number 5 is an advertisement text type with a synonym question type.
6. Item number 6 is a label text type with a purpose of the text question type.
7. Item number 7 is a label text type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
8. Item number 8 is a label text type with a synonym question type.
9. Item number 9 is a procedure text type with a topic of the text question type. The topic of the text is what is discussed in the text as a whole.
10. Item number 10 is a procedure text type with general inference question types. Inference questions are questions that are required to use clues in the text to conclude what the text is talking about, in other words the questions are implicit.
11. Item number 11 is a procedure text type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.

12. Item number 12 is a narrative text type with a topic of the text question type.
The topic of the text is what is discussed in the text as a whole.
13. Item number 13 is a narrative text type with factual information question types.
Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
14. Item number 14 is a narrative text type with the type of inference question but more specifically what is sought is moral value.
15. Item number 15 is a procedure text type with a purpose of the text question type.
16. Item number 16 is a procedure text type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
17. Item number 17 is a procedure text type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
18. Item number 18 is a procedure text type with a reference question type.
Reference or commonly referred to as refers to is a reference to provide information or support for a text/subject/object.
19. Item number 19 is a report text type with a topic of the text question type. The topic of the text is what is discussed in the text as a whole.
20. Item number 20 is a report text type with factual information question types.
Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.

21. Item number 21 is a report text type with factual information question types. Factual information is a question that asks for information that is clearly written, in other words it is explicit or certain information.
22. Item number 22 is a report text type with the type of inference question but more specifically the target is the audience.
23. Item number 23 is a report text type with a purpose of the text question type.
24. Item number 24 is a report text type with main idea question type. Main idea is a question about finding the main idea of a paragraph.
25. Item number 25 is a report text type with general inference question types. Inference questions are questions that are required to use clues in the text to conclude what the text is talking about, in other words the questions are implicit.

b. The Reliability of the English Summative Test

This reliability works on the standard index described by Arikunto (2013) which highlights that an item is said to be reliable if the Cronbach's Alpha value obtained is at least 0.60 (p. 221).

Based on the results of interviews with school principals, English teachers and students, the summative test for English subjects is usually 50 multiple choice items, but due to the Covid-19 pandemic, the curriculum was narrowed based on the agreement of English teachers and in accordance with the approval of the Education and Culture Office of Hulu Sungai Utara Regency, so the summative test for English subjects only become 25 multiple choice items with a processing time of 90 minutes.

According to the English teacher at SMPN 3 Sungai Pandan, the distribution of the items carried out by the ninth grade students of SMPN 3 Sungai Pandan was also tested in all SMPN in Hulu Sungai Utara with assurance that there would be no question leaks. based on the statement given by the teacher, the level of reliability of the items carried out by students, the estimated reliability coefficient on the items increased. The items that are made are also based on the competencies listed in the 2013 curriculum by paying attention to the indicators in each syllabus item.

In interviews conducted with several students via Whatsapp, the level of difficulty in the items presented was in accordance with the learning provided. Some students have difficulty answering the questions because the students themselves do not pay attention and repeat the lessons that have been given by the teacher.

On the other hand, during the online English learning process, most of the students feel that the teaching and learning process is not effective. the government made a regulation that all schools in Indonesia had to hold virtual learning and every academic activity was carried out online due to the pandemic situation. The Researcher understand the situation due to the pandemic and students also enter a certain age to advance to high school so that the final exam must continue. Actually, the researcher knows that the students' truth can't be relied on, because they can do the final exam with their friends or through chat, or maybe video calls, and group messengers. However, looking at the data on student scores, only a few were above

limit scored, so the researcher speculated that no one would help them in doing the summative test items.

2. Quantitative Phase

a. The Validity of the English Summative Test

Quantitative data was taken from the answer sheets of students of class XI SMPN 3 Sungai Pandan as many as 65 answer sheets. The researcher has identified the items given to students and have counted the number of items which are valid and invalid. The researcher compiled the results of calculations using the SPSS 26.0 program. The findings data show that twenty-two items of the English summative test are valid and three items are invalid. For more details, the researcher provides a table that provides a brief description of the validity of each item.

Item	r count	r table	Status
1	0.451	0.2441	Valid
2	0.505	0.2441	Valid
3	0.450	0.2441	Valid
4	0.447	0.2441	Valid
5	0.371	0.2441	Valid
6	0.461	0.2441	Valid
7	0.225	0.2441	Invalid
8	0.231	0.2441	Invalid

9	0.357	0.2441	Valid
10	0.402	0.2441	Valid
11	0.321	0.2441	Valid
12	0.610	0.2441	Valid
13	0.616	0.2441	Valid
14	0.417	0.2441	Valid
15	0.433	0.2441	Valid
16	0.597	0.2441	Valid
17	0.364	0.2441	Valid
18	0.347	0.2441	Valid
19	0.183	0.2441	Invalid
20	0.339	0.2441	Valid
21	0.539	0.2441	Valid
22	0.488	0.2441	Valid
23	0.462	0.2441	Valid
24	0.316	0.2441	Valid
25	0.255	0.2441	Valid

There are four columns in the table, the first column provides information about the test item number. Then, the second column provides information about the results of the validity analysis or r count. Next, the third column provides information about the critical value of the product moment table with a significance

level of 95% or r table. The last, fourth column provides information about the validity status.

From the table above, it can be seen that the valid items are item numbers 1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24 and 25. Then, the invalid items are items 7, 8, and 19. For more details, the researcher describes each item as follows;

1. Item number 1 is a valid item because the result of r count is higher than the product moment table or r table.
2. Item number 2 is a valid item because the result of r count is higher than the product moment table or r table.
3. Item number 3 is a valid item because the result of r count is higher than the product moment table or r table.
4. Item number 4 is a valid item because the result of r count is higher than the product moment table or r table.
5. Item number 5 is a valid item because the result of r count is higher than the product moment table or r table.
6. Item number 6 is a valid item because the result of r count is higher than the product moment table or r table.
7. Item number 7 is an invalid item because the result of r count is lower than the product moment table or r table.
8. Item number 8 is an invalid item because the result of r count is lower than the product moment table or r table.

9. Item number 9 is a valid item because the result of r count is higher than the product moment table or r table.
10. Item number 10 is a valid item because the result of r count is higher than the product moment table or r table.
11. Item number 11 is a valid item because the result of r count is higher than the product moment table or r table.
12. Item number 12 is a valid item because the result of r count is higher than the product moment table or r table.
13. Item number 13 is a valid item because the result of r count is higher than the product moment table or r table.
14. Item number 14 is a valid item because the result of r count is higher than the product moment table or r table.
15. Item number 15 is a valid item because the result of r count is higher than the product moment table or r table.
16. Item number 16 is a valid item because the result of r count is higher than the product moment table or r table.
17. Item number 17 is a valid item because the result of r count is higher than the product moment table or r table.
18. Item number 18 is a valid item because the result of r count is higher than the product moment table or r table.
19. Item number 19 is an invalid item because the result of r count is lower than the product moment table or r table.

20. Item number 20 is a valid item because the result of r count is higher than the product moment table or r table.
21. Item number 21 is a valid item because the result of r count is higher than the product moment table or r table.
22. Item number 22 is a valid item because the result of r count is higher than the product moment table or r table.
23. Item number 23 is a valid item because the result of r count is higher than the product moment table or r table.
24. Item number 24 is a valid item because the result of r count is higher than the product moment table or r table.
25. Item number 25 is a valid item because the result of r count is higher than the product moment table or r table.

b. The Reliability of the English Summative Test

The data from the research show that the summative test of English class IX students of SMPN 3 Sungai Pandan is reliable because the reliability index is 0.785. For more details, the researcher provides a table of reliability analysis as follows.

Table 4.2		
The Reliability Analysis		
Cronbach's Alpha	r table	Status
0.785	0.60	Reliable

There are three columns in the table, the first column provides information about the Cronbach's Alpha values found by the researcher. Then, the second column provides information about the minimum value of Cronbach's Alpha for reliable items. And the last, third column provides information about the reliability status.

B. Discussion

This study aims to identify the validity and reliability of the summative test items for grade IX students at SMPN 3 Sungai Pandan, Sungai Pandan District, Hulu Sungai Utara Regency. Based on the research data analyzed with the help of the SPSS version 26.0 software program, this section will discuss the results of the research that has been carried out.

1. The Validity of the English Summative Test

Based on the findings, the results of the existing test data report that twenty-two test items are valid and three items are invalid. This fact simply gives us an idea of the current state of the English summative test used for grade IX students at SMPN 3 Sungai Pandan.

Validity test is used to determine the validity of the instrument research, meaning to find out whether the instrument has been able to measure what is intended to be measured. Instrument validity is one of the characteristics that mark an instrument as good. In this study, researchers used the product moment correlation formula from Karl Pearson (Suharsimi, 2010, p. 213). In line with this, according to Sugiyono (2010), "the decision-making criteria in determining the

validity of the instrument were consulted directly on the r product moment table” (p. 455). Therefore, invalid items need to be removed or revised and activities must be carried out by the teacher to match the normal validity index of a high-quality test. This information should enable the test constructor to master item validity analysis with the aim of creating test items that work on the item’s ability to measure what it is supposed to measure.

From 25 English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021 tested. There are 22 items that are valid and can be said already good. Those are number 1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, and 30 with an analysis as below:

1. No. 1. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power are good, but effectiveness of distractors is should be revised or changed. Since, option B no student chooses. Preferably, the word on option B “to encourage” can be replaced with the word “to motivate” so that all distractors function. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
2. No. 2. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.

3. No. 3. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
4. No. 4. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
5. No. 5. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
6. No. 6. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
7. No. 9. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next

time and clarified by the results of the validity test which shows that this item is valid.

8. No. 10. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
9. No. 11. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
10. No. 12. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
11. No. 13. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.

12. No. 14. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
13. No. 15. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
14. No. 16. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
15. No. 17. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
16. No. 18. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used

in the next time and clarified by the results of the validity test which shows that this item is valid.

17. No. 20. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
18. No. 21. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid
19. No. 22. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power are good, but effectiveness of distractors is should be revised or changed. Since, option D no student chooses. Preferably, in option D be replaced with a more parallel word, from the word “A historian” to “A journalist” and more possible to be selected by students so that all distractors function. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
20. No. 23. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used

in the next time and clarified by the results of the validity test which shows that this item is valid.

21. No. 24. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.
22. No. 25. The test item is good. So, the item does not need to be revised or changed because based on the stem, the level of difficulty, discriminatory power, and effectiveness of distractors are good. Overall, this item can be used in the next time and clarified by the results of the validity test which shows that this item is valid.

From 25 English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021 tested. There are 3 items that are not valid. Those are number 7, 8, and 19. The analysis is as follows:

1. No. 7

The product is in the form of

- | | | |
|-------------|---------------------------|-------------|
| A. liquid | | A. pill |
| B. tablets | It should be revised
→ | B. tablets |
| C. powder | | C. caplets |
| D. capsules | | D. capsules |

The item should be revised or changed because the difficulty level is too easy where almost all students can answer correctly. Since, the answer to this item

is too clear so students can easily answer it, so the answer is very easy for students to guess. This item cannot be used in the next time because too easy and the answer too clear.

The distractors must be revised or replaced because they are not parallel, especially for options A and C. Since, they are not effective which most students ignore. It is recommended that these two distractors be replaced, such as distractor A from the word “Liquid” being replaced with the word “Pill” and distractor C from the word “Powder” being replaced by the word “Caplets”. This distractor cannot be used later because it is non-parallel and non-specific.

2. No. 8

“.. if you are pregnant or lactating consult a healthcare professional before ...”

The underlined phrase is synonymous with

A. a psycho

A. a nurse

B. a psychic

It should be revised

B. a midwife

C. a physician

C. a doctor

D. a psychologist

D. a pharmacist

The item should be revised or changed because the difficulty level was too difficult where more than half of the students were unable to answer correctly. The supporting factor why this item invalid is the power of discrimination is poor. The item cannot distinguish between the upper and lower students because all options are almost the same even though they have different meanings and it is possible for students to answer carelessly.

Preferably, use options with words that are more familiar to students, such as option A “a nurse”, option B “a midwife”, option C “a doctor”, and option D “a pharmacist”. This item cannot be used at a later time, as choice and distraction are too difficult and allows students to answer carelessly due to confusion and is proven by two students who did not answer this item.

3. No. 19

What is the text mostly about?

- A. The formation and the shape of tornado
- B. The effect and the size of tornado
- C. The cause and the danger of tornado
- D. The kinds and the effect of tornado

It should be revised ↓

What is the text general about?

- A. The origin and history of tornado
- B. The effect and the size of tornado
- C. Types of cumulonimbus and the meaning of tornado
- D. The impact and frequently affected areas of a tornado

The item should be revised or changed because the level of difficulty is too difficult where most of the students cannot answer it correctly. Since, stem sentences are not common in spoken language. Preferably, the word “mostly” is replaced with the word “general” to make it easier for students to understand.

The power of discrimination is bad. The item cannot distinguish between upper and lower students. Since, the level of difficulty that can be said is very difficult and allows students to answer carelessly because of the lack of clarity between the options and many students are distracted by option A. It is better if the options are taken from the sentences in the text so that students are able to distinguish each option. For example, option A “The origin and history of tornado”, option B “The effect and the size of tornado”, option C “Types of cumulonimbus and the meaning of tornado”, and option D “The impact and frequently affected areas of a tornado”. The distractor should be revised or changed for all options except the answer key, because it is not effective where most of the students are distracted. This item cannot be used later because the questions are very difficult, the discriminatory power and distractors are bad.

2. The Reliability of the English Summative Test

Based on the findings, the results of the data show that the summative test of English for grade IX students of SMPN 3 Sungai Pandan with a total of 25 items is reliable because the index is more than 0.60, which is 0.785. This fact illustrates that the condition of the English summative test used for grade IX students at SMPN 3 Sungai Pandan is currently also reliable.

According to Arikunto (2016) as one part of the test requirements, reliability has a relationship with trust issues. A test is said to be reliable, if the test gives a constant result even though it is given many times. In other words, the results show consistency or determination (p. 100). The instrument is said to be reliable if the

Cronbach's Alpha value is > 0.600 , so if the Cronbach's Alpha value is less than 0.600 the instrument is said to be unreliable. Added by an expert opinion, namely the second main requirement of the data collection instrument is reliability. The instrument is said to be reliable if the instrument is used several times to measure the same object, it will produce the same data (Sugiyono, 2010, p. 348). In other words, although there are some items out of 25 invalid items based on the validity test taken from a population of 65 students, the summative test items of English class IX at SMPN 3 Sungai Pandan can still be said to be reliable because Cronbach's Alpha shows a score that exceeds the minimum limit.

Reliability relates to the ability of measuring instruments to make measurements accurately. Reliability is the accuracy and precision produced by measuring instruments in making measurements. Empirically, the level of reliability is indicated by a number called the reliability coefficient. The empirical reliability of objective questions is divided into three types, namely the stability coefficient, the equivalent coefficient, and the coefficient of internal consistency. Factors that affect reliability, namely the length of the test, the distribution of questions, the difficulty of the test, and objectivity.

In these items, the reliability of the achievement to be achieved by the items as a measuring instrument can be said to be successful, although there are some items that have not been effective, but most of the items are good in terms of reliability. The duration of the test is considered sufficient to complete 25 item multiple choice in 90 minutes, although there are some children who have not been able to complete all of them. Even in the distribution of items, students are only

given a grid for them to study and repeat the material to be tested so that the level of difficulty of this item is relatively easy and can be reached by students and can be trusted and relied on.