

A THESIS

**AN ANALYSIS OF THE ENGLISH TEACHER-MADE TEST FOR
SECOND GRADE STUDENTS AT UPT SMKN 3 PAREPARE**



BY

RISMA
Reg. Num. 19.1300.078

**ENGLISH EDUCATION PROGRAM
TARBIYAH FACULTY
STATE ISLAMIC INSTITUTE
PAREPARE**

2023

A THESIS

**AN ANALYSIS OF THE ENGLISH TEACHER-MADE TEST FOR
SECOND GRADE STUDENTS AT UPT SMKN 3 PAREPARE**



BY

**RISMA
Reg. Num. 19.1300.078**

Submitted to the English Education Program of Tarbiyah Faculty of State Islamic
Institute of Parepare in Partial of Fulfilment of the Requirements
For the Degree of Sarjana Pendidikan (S.Pd.)

**ENGLISH EDUCATION PROGRAM
TARBIYAH FACULTY
STATE ISLAMIC INSTITUTE
PAREPARE**

2023

ENDORSEMENT OF CONSULTANT COMMISSIONS

The Title of Thesis : An Analysis of the English Teacher-Made Test
For Second Grade Students at UPT SMKN 3
Parepare

Name of the Student : Risma

Student Reg. Number : 19.1300.078

Study Program : English Education Program

Faculty : Tarbiyah

By Virtue of Consultant Degree : The Dean of Tarbiyah Faculty
No. 3223 Tahun 2022



Approved by
Consultant Commissions

Consultant : Dr. Magdahalena Tjalla, M.Hum.

NIP : 19700320 200501 2 006

Co-Consultant : Yessicka Noviasmy, M.Pd.


NIDN : 2029119004

Cognizant of:

Tarbiyah Faculty
Dean




 Dr. Zulfah, M.Pd.
 NIP. 19830420 200801 2 010

ENDORSEMENT OF EXAMINER COMMISSIONS

The Title of Thesis : An Analysis of the English Teacher-Made
Test For Second Grade Students at UPT
SMKN 3 Parepare

Name of the Student : Risma

Student Reg. Number : 19.1300.078

Faculty : Tarbiyah

Study Program : English Education Program

By Virtue of Consultant Degree : The Dean of Tarbiyah Faculty
No. 3223 Tahun 2022

Date of Graduated : July 24th 2023

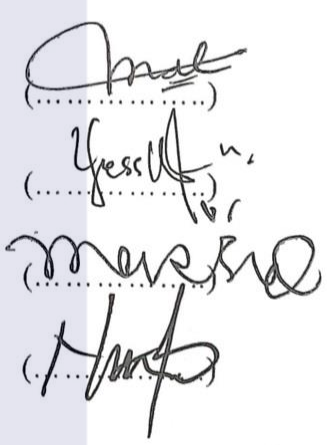
Approved by Examiner Commissions

Dr. Magdahalena Tjalla, M.Hum. (Chairman)

Yessicka Noviasmy, M.Pd. (Secretary)

Dr. Arqam, M.Pd. (Member)


Nur Asiza, M.Pd. (Member)



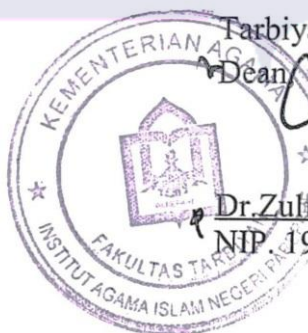
Cognizant of:

Tarbiyah Faculty

Dean


Dr. Zulfah, M.Pd.

NIP. 19830420 200801 2 010



ACKNOWLEDGEMENTS

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
 الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ وَالصَّلَاةُ وَالسَّلَامُ عَلَى أَشْرَفِ الْأَنْبِيَاءِ
 وَ الْمُرْسَلِينَ وَعَلَى آلِهِ وَصَحْبِهِ أَجْمَعِينَ أَمَّا بَعْدُ

Alhamdulillah Rabbil Alamin. First of all, the writer would like to express her grateful to the lord Allah swt. The highest and the master of the universe while the creator of the everything in this world who still lend the writer good health, blessing, mercy him can finish this thesis. Salam and Salawat always are sent to the prophet Muhammad saw. Peace be upon him. Who has guided us from uneducated person to be educated person.

This wonderful time, the researcher wants to thank a lot to all people who have supported and helped her in finishing her skripsi as effective as she can. She realizes that without their support and help, she could not be able to finish this “thesis”. This opportunity, the writer would like to express very especial thanks to beloved parents, Isa and Usman who have given an endless love, advice, support and praying to be success students ever. Her high appreciation and profusely sincere thanks are due to Dr. Magdahalena Tjalla, M.Hum as the first consultant and Yessicka Noviasmy, M.Pd. as the second consultant who have patiently guided and given their construction suggestion, motivation and correction to the writer for finishing this thesis. Another side, the writer would also deliver special thanks to:

1. Dr. Hannani, M.Ag., as the Rector of IAIN Parepare who has worked hard to manage education at IAIN Parepare.

2. Dr. Zulfah, S.Pd, M.Pd as Dean of Tarbiyah Faculty IAIN Parepare, who has arranged a positive education for the students in the Faculty of Tarbiyah and also allowed the writer to research in this Tarbiyah Faculty.
3. Dr. Arqam, M.Pd as The Chairman of English Program for the fabulous serving to the students.
4. All lecturers of English Program who have already taught the writer during her study in IAIN Parepare.
5. St. Ruwaedah, S.Pd. as the Teacher of the XI Multimedia and thank you for your help. Specially, students in XI Multimedia 1 the class where the researcher did the research, they are so amazing.
6. Great thanks to best friend Umailiana Firra Rahma, Azra Ahmad Wardi, Mutmainna, and all family Btn Graha.
7. All people who have given their help in writing this thesis that the writer could not mention it one by one.

May Allah SWT always bless them. The researcher realizes that the Skripsi is far from perfect and need to be complete for being perfect. Therefore, the suggestions will be accepted to make this thesis is better.

Parepare, May 27th 2023
30 Syawal 1444 H
The Researcher,



Risma
Reg. Num. 19.1300.078

DECLARATION OF THE AUTHENTICITY OF THE THESIS

The researcher who signed the declaration below:

Name : Risma
Student Reg. Number : 19.1300.078
Place and Date of birth : Mamuju, 08th December 2000
Study Program : English Education
Faculty : Tarbiyah Faculty
Tittle of Skripsi : An Analysis of English Teacher-Made Test For
Second Grade Students at UPT SMKN 3 Parepare

Stated this thesis was her own writing and if it can be proved that it was copied, duplicated or complied by any other people, this thesis and the degree that has been gotten would be postponed.

Parepare, May 27th 2023

The Researcher,



Risma
Reg. Num. 19.1300.078

ABSTRACT

Risma. *An Analysis of the English Teacher-Made Test For Second Grade Students at UPT SMKN 3 Parepare* (Supervised by Magdahalena and Yessicka Noviasmy)

This study aims to determine the quality of the multiple-choice final test items for class XI Multimedia 1 at UPT SMKN 3 Parepare for the 2022-2023 academic years in terms of validity, reliability, level of difficulty, discrimination index, and distractors.

The type of research used in this research was quantitative descriptive. The subjects in this study were 33 students in class XI Multimedia 1. The data in this study was collected used documentation techniques.

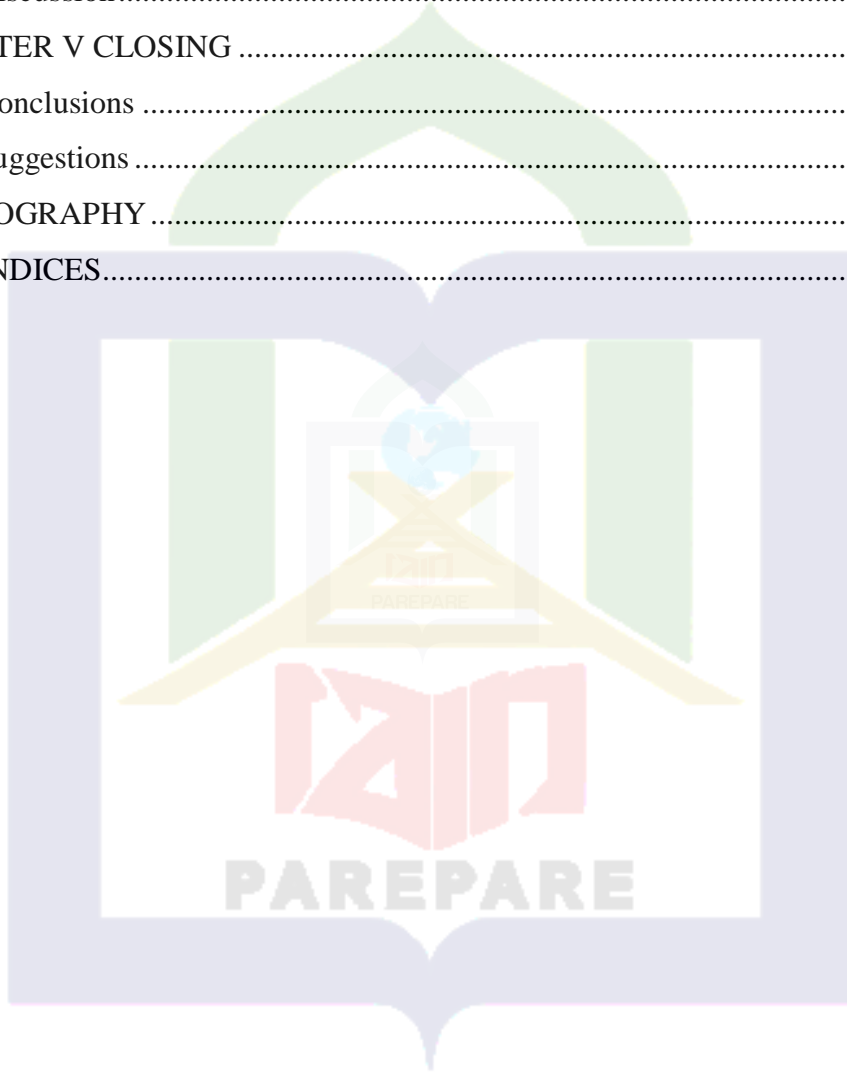
The results of the study showed that: Valid category is 8 items and invalid are 2 items. Included in the test that is unreliable. Difficult category is 1 item, moderate category is 8 items, and easy category is 1 item. Worst discrimination index category is 1 item, satisfactory category is 1 item, good category is 6 items and excellent category is 2 items. Distractors function very good is 7 items, good category are 2 items, and quite good category is 10 items. It can be concluded that the final multiple-choice test for class XI Multimedia 1 at UPT SMKN 3 Parepare has very good quality of item validity, unreliable test, good discrimination index, good quality level of difficulty and very good quality distractor.

Keywords: Item Analysis, Teacher-Made Test

TABLE OF CONTENTS

COVER	i
ENDORSEMENT OF CONSULTANT COMMISSIONS	ii
ENDORSEMENT OF EXAMINER COMMISSIONS	iii
ACKNOWLEDGEMENTS	ii
DECLARATION OF THE AUTHENTICITY OF THE THESIS	vi
ABSTRACT	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	x
LIST OF APPENDICES	xi
CHAPTER I INTRODUCTION	1
A. Background	1
B. Research Questions	4
C. Objective of The research	4
D. Significant of The Research	4
CHAPTER II REVIEW OF RELATED LITERATURE	6
A. Previous Related Research Findings	6
B. Some Pertinent Ideas	10
C. Theoretical Framework	22
D. Hypothesis	23
CHAPTER III METHODOLOGY OF RESEARCH	24
A. The Research Design	24
B. Location and Time	24
C. Population and Sample	25
D. Technique of Data Analysis	26
E. Instrument of the research	26

F. Operational Definition of Variables.....	26
G. Technique of Data Analysis.....	28
CHAPTER IV FINDINGS AND DISCUSSIONS	32
A. Findings	32
B. Discussion.....	39
CHAPTER V CLOSING	49
A. Conclusions	49
B. Suggestions	50
BIBLIOGRAPHY	I
APPENDICES.....	IV



LIST OF TABLES

Table of Number	The Tittle of Tables	Pages
3.1	Criteria used in analyzing discrimination index	30
4.1	Student response score XI Multimedia 1 on the final test multiple-choice at UPT SMKN 3 Parepare	33
4.2	The result of an English Teacher-Made Test Analysis Based on the Validity.	34
4.3	The result of an english teacher-made test analysis based on the level of difficulty.	36
4.4	The result of an english teacher-made test analysis based on discrimination index	37
4.5	The number of options selected by students	38
4.6	The result of an english teacher-made tests analysis based on distractors	39

LIST OF APPENDICES

Number of Appendix	The Title of Appendices
1.	Title Certificate
2.	Instruments of The Research
3.	Research Papers
4.	Final Test For Second Grade at UPT SMKN 3 Parepare
5.	The Result Calculation For Test Students



CHAPTER I

INTRODUCTION

A. Background

A test is a tool having a set of questions and an examination. A test is used to measure a particular characteristic of an individual or group of individuals. A test is designed to measure the quality, ability, skill, or knowledge of a sample against a given standard, which can usually be deemed as acceptable or not. Manichander stated that although the word "test" has been interchangeably used to mean assessment or even evaluation, the distinguishing factor of a test is the fact that it is a form of assessment.¹

Brown said that tests are divided into two forms, namely teacher-made tests and standardized tests.² A teacher-made test is designed by the teacher based on the curriculum and the lesson plan that have been applied during the lesson. It is intended to measure the success rate of students in achieving the target of the curriculum after the teaching learning process is done by the teacher. Therefore, the teacher must make logical and rational questions about what items are worth asking. This test is usually used for daily, formative, and general tests.³

Test is a way of measuring and assessing in the field of education in the form of giving assignments or a series of tasks that must be done by students, so that

¹ Dickson Adom, Jephtar Adu Mensah, and Dennis Atsu Dake, "Test , Measurement , and Evaluation: Understanding and Use of the Concepts in Education" 9, no. 1 (2020): 109–19, <https://doi.org/10.11591/ijere.v9i1.20457>.

² H.Douglas Brown, *Language Assessment Principle and Classroom Practices* (New York: Pearson Education Inc, 2004).

³ Lukmanul Hakim and Irhamsyah, "The Analysis Teacher-Made Test for Senior High School at State Senior High School 1 Kutacane" 21, no. 1 (2020): 10–20.

students' achievement scores can be known.⁴ The test results obtained from students will be an indication of whether the test used is good or not. The quality of a test can be seen from the characteristics of the questions used, a test with good quality will have good items. To find out whether the questions have good quality, each item must analyze.

An analysis is a systematic procedure that will provide very specific information on something very detailed, such as an item analysis of what has been compiled. Item analysis is an activity that the teacher must do to improve the quality of what has been written. According to Aiken, the purpose of item analysis is to improve the quality of item tests and find out student diagnostic information. Quality questions are questions that can provide information as accurately as possible, so that it can be seen that they have mastered the material and those who have not.

The teacher has the obligation to make semester exam questions, and the questions that are made must be proportionally appropriate. The items have good and functioning instruments, meaning that with these instruments, the measuring instruments used actually contain the material to be measured and are in accordance with the compilers of the questions. Therefore, teachers must really pay attention to the quality that will be tested on students in order to obtain the correct results so that the quality of students can be guaranteed.

In fact, teacher rarely does trials on the questions to be used, including analysing the quality of each item to be tested, so most of them cannot identify good or bad tests. This is due to the lack of teacher time and teacher understanding. Even

⁴ Yohana Fransisca Dewi Risna Putri, "An Analysis of Teacher-Made English Final Second Semester Test for the Year Eleven Students Of SMAN 1 Ambarawa In The Academic Year Of 2008/2009 Based on the Representativeness of Content Standard" (Semarang State University, 2009).

though analyzing the items is an activity that must be carried out by a teacher. According to Nurgiyantoro, the reason for doing item analysis is because it will produce quality questions on subsequent tests and to find out the strengths and weaknesses of items so that items can be selected, revised, and the problem with the item found immediately.⁵

If the tests made by the teacher are not in accordance with what has been determined, it will have an adverse effect on students; namely, the interests, talents, and understanding of students cannot be measured, so the teacher cannot classify students who have high abilities and students who have low abilities.

According to the findings of an interview with an English teacher at UPT SMKN 3 Parepare named St. Ruwaedah, she never does a test analysis of the final test that has been based on validity, reliability, level of difficulty, discrimination index, and distractors. In addition, the teacher also does not really understand the problem of the test instrument, and the teacher does not have time to analyze the instrument test that has been made to determine whether the exam questions that have been made meet the standards, terms, and conditions contained in the assessment tool.

UPT SMKN 3 Parepare is a school with very good accreditation. This school is located at Jl. Karaeng Burane No. 16, Parepare, and South Sulawesi. UPT SMKN 3 Parepare has six departments or expertise competencies, namely: travel business, catering, hair beauty, fashion, multimedia, and broadcasting techniques for the production and broadcasting of television programs.

This research is very important because it is to find out the quality of the test made by the teacher. The quality of the test in question is whether the test is valid,

⁵ Nurgiyantoro B, *Penilaian Pembelajaran Bahasa* (Yogyakarta: BPFE_Yogyakarta, 2010).

reliable, the questions are of good quality, the questions are easy to answer or difficult to answer, and the distractor functions properly. In addition, this research is important so that teachers know how to analyze multiple-choice test. And the hope after this research is done is that the teacher is able to analyze the tests made and other tests for his students.

Based on the description above, it is necessary to analyze the items in detail quantitatively to test results for validity, reliability, level of difficulty, discrimination index, and distractors. Researcher interested in conducting research with the title "An Analysis of the English Teacher-made Test for Second Grade Students at UPT SMKN 3 Parepare".

B. Research Questions

Based on the background above, the researcher give formula the problem of the research as bellows:

How is the quality of teacher-made test based on validity, reliability, level of difficulty, discrimination index, and distractors?

C. Objective of The research

Based on the researcher formulate the problem previously set out the report of this researcher aimed at finding out the following objective:

To know the quality of teacher-made test based on validity, reliability, level of difficulty, discrimination index, and distractors.

D. Significant of The Research

1. The results of this study are expected to contribute to the world of education as a reference for teachers, especially item analysis activities so that they can improve the quality of education.

2. Provide information to teachers and prospective teachers to help them pay attention to the quality of the test instruments that will be used in learning.
3. This research is expected to be used as material for conducting further research and developing the process of item analysis on other subjects.



CHAPTER II

REVIEW OF RELATED LITERATURE

A. Previous Related Research Findings

There are many research findings which are related to this research, some of previous research findings which are related to this research are described below:

The first research was written by Neti Hartati and Hendro Pratama Supra Yogi which tested the quality of multiple choice items in terms of difficulty level, discriminating power, and the effectiveness of distractors. The summative English test and the students' answer sheets were utilized as the basis for the qualitative method employed in this study. The English Summative Test's item quality is described used the qualitative analysis. 65 students are the study's subject. The results showed that, while there should only be 12–13 items, the number of simple items is larger than anticipated (19 or 38%). Some things' discrimination index levels are low, and they are rejected with low discriminating power. Further investigation reveals that the two issues are caused by the significant amount of poor or rejected distractors.⁶

The similarity of research conducted by prospective researchers is the focus of research on the quality of test made by English teacher. But Neti and Hendro's research used qualitative and quantitative research, while the research conducted by the prospective researcher used quantitative descriptive. Neti and Hendro's research does not used validity and reliability, while the research conducted by the prospective

⁶ Neti Hartati and Hendro Pratama Supra Yogi, "Item Analysis for a Better Quality Test," 2003, 12.

researcher used validity, reliability and distractors to know the quality of teacher-made test.

The second research, written by Mutiara Kusumawati and Samsul Hadi, was carried out to statistically assess the standard of the final Mathematics examination items. Descriptive quantitative research was used in this study. The 353 student sample drawn used the purposive sampling technique provided the data. This study revealed that there were no easy items and that 40% of the 35 items assessed were extremely challenging. The item discrimination index is described as follows: A high item discrimination index value is held by 5.71% of the items, a very high item discrimination index value by 2.86% of the items, a medium item discrimination index value by 31.43% of the items, and a very low item discrimination index value by 8.57% of the items.⁷

Mutiara Kusumawati and Samsul Hadi in her research entitled "An Analysis of Multiple Choice Questions (MCQs): Item and Test Statistics from Mathematics Assessments in Senior High School". The author used final examination Mathematic subject for class X and used Item Response Theory (ITR). The authors do not use validity and reliability, while the research conducted by the prospective researcher used validity and reliability. The similarity of research conducted by prospective researchers is the focus of research on the quality of tests made by teacher and used quantitative research.

Maya Marsevani conducted a study aimed at determining the quality of multiple-choice questions in a public elementary school in terms of difficulty level,

⁷ Mutiara Kusumawati and Samsul Hadi, "An Analysis of Multiple Choice Questions (MCQs): Item and Test Statistics from Mathematics Assessments in Senior High School" 4, no. 1 (2018): 70–78.

discrimination power, and distractor. To gather data and assess MCQs on student assessments, this study used a cross-sectional survey. 40 students are the study's subject. This study demonstrated that the item discrimination power was strong and that the majority of the difficulty index was acceptable. Two objects were entirely effective as a distractor.⁸

Maya's research employed a cross-sectional study to obtain information and evaluate multiple-choice questions in the students' tests, while the research conducted by the prospective researcher used documentation and quantitative descriptive. Maya's research does not use validity and reliability, while the research conducted by the prospective researcher used validity and reliability. The similarity of research conducted by prospective researchers is the focus of research on the quality of the English teacher-made test.

Lukmanul Hakim and Irhamsyah conducted study to find out whether or not the English test made by teacher was valid. This study used qualitative descriptive research. The description is based on a test the instructor created or a final test question sheet she used with her Senior High School 1 Kutance students. It was a documentary analysis technique that was used. This study demonstrated that the Senior High School 1 Kutance English teacher had developed legitimate test questions, and the degree of validity of those questions demonstrated that teachers are capable of developing legitimate English examination questions.⁹

The difference between the research conducted by prospective researchers and Lukmanul Hakim and Irhamsyah's research is that prospective researchers used

⁸ Maya Marsevani, "Item Analysis of Multiple-Choice Questions : An Assessment of Young Learners" 10, no. 2 (2022): 401–8.

⁹ Lukmanul Hakim and Irhamsyah, "The Analysis Teacher-Made Test for Senior High School at State Senior High School 1 Kutacane" 21, no. 1 (2020): 10–20.

quantitative descriptive research, whereas Lukmanul Hakim and Irhamsyah's research used qualitative descriptive research, and Lukmanul Hakim and Irhamsyah's research only used validity, whereas prospective researchers used validity, reliability, level of difficulty, discrimination index, and distractors to know how the quality of teacher-made tests is. The similarity of research conducted by prospective researchers is the focus of research on the quality of the English teacher-made test.

Taufiq Effendi and Ilza Mayuni conducted study to examine the quality of a multiple-choice, teacher-made English formative informal assessment for four classes of high school students of an English course in Madura. In this study, an item response analysis was performed. Item Response Theory (IRT) serves as the foundation for such analysis. Data were gathered from the scores of 80 students who took the test. In order to ensure that pre-service teachers and in-service teachers have the essential test development and test analysis skills for continuous improvement of the learning, teaching, and assessment practices, the findings advise that schools and teacher institutes give the relevant trainings.¹⁰

The difference between Taufiq Effendi and Ilza Mayuni's research and the research conducted by prospective researchers was that Taufiq Effendi and Ilza Mayuni's research do not use level of difficulty, whereas prospective researcher used level of difficulty to know how the quality of teacher-made tests is. And prospective researchers used quantitative descriptive research, whereas Lukmanul Hakim and Irhamsyah's research used an Item Response Theory (IRT). The similarity of research conducted by prospective researchers is the focus of research on the quality of the English teacher-made test.

¹⁰ Taufiq Effendi and Ilza Mayuni, "Examining Teacher-Made English Test in a Language School" 2, no. 2 (2022): 67–76.

The novelty in this research is to analysis of the English Teacher-Made test for second grade students at UPT SMKN 3 Parepare based on 5 aspects namely: validity, reliability, level of difficulty, discrimination index, and distractors.

From some of the research results that have been described above, there are quite clear differences between previous research and the research that carried out by researchers. Therefore, this time the researcher tries to examine An Analysis of the English Teacher-Made Test for Second Grade Students at UPT SMKN 3 Parepare based on validity, reliability, level of difficulty, discrimination index, and distractors.

B. Some Pertinent Ideas

1. Definition of Test

Test is a way to measure the knowledge ability of person or student. According to Brown, test is a method of measuring a person's abilities, knowledge or performance.¹¹ The test is a number question that has right or wrong answers. Test is an assessment tool in the form of a task that students must do to get an objective and accurate score.¹²

Suharsimi Arikunto reveals that:

A test is a collection of several questions or other tools used to measure skills, intellectual knowledge, abilities or talents possessed by individuals or group.

Rosnita explained in more detail that a test is a formal instrument used to assess students' cognitive abilities in a subject as well as to collect quantitative information about students' psychomotor abilities (physical skills) and affective characteristics (such as attitudes, emotions, interests, and values). A test generally

¹¹ H.Douglas Brown, *Language Assessment Principle and Classroom Practices* (New York: Pearson Education Inc, 2004).

¹² Kurniawan Haryanto, "Analisis Butir Soal Pilihan Ganda Ulangan Akhir Semester Genap" (Universitas Sanata Dharma, 2016).

includes a series of questions, statements, or tasks that are arranged for a student or group of students.¹³

Based on some of the above, it can be concluded that the test is measuring tool in the form of tasks or questions that must be answered by students' which used as a means of collecting information on behavior or learning achievement results of participants during the implementation of learning activities. Test is a method in determining students' ability to complete certain tasks.

2. Types of test

Since the researcher wants to conduct an analysis of the English teacher-made test, the researcher will present some information distinguishing the teacher-made test of classroom from those of standardized testing and the difference between formative and summative assessment.

a. Teacher-made test

According to Suharsimi Arikunto, a teacher-made-test is a test written and made by a teacher in the school, so the validity and reliability of the test is not like a standardized test.¹⁴ The effectiveness of this type of test depends on the skill and ability of the teacher in designing the test. This test is based on materials and specific goals formulated by the teacher for his own class. Rarely does the teacher analyze and revise test items that have been tried, so the teacher does not know the validity, reliability, level of difficulty, discrimination index, and distractors.

¹³ Asrul, Rusydi Ananda, and Rosnita, *Evaluasi Pembelajaran* (Medan: Citapustaka Media, 2014).

¹⁴ Suharsimi, Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, 2nd ed. (Jakarta: Bumi Aksara, 2013).

Brown mentioned that, basically, the teacher-made test can only be used in some classes the teacher teaches.¹⁵ The advantage of using this kind of test is that students are familiar with the task given by the teacher, which enables them to score better. Meanwhile, the limitations of teacher-made test are limited sampling, low reliability, subjective, low validity, high skill required, monotonous, and time-consuming.

b. Summative and Formative Assessment

Summative test contain task that must be done by students after the learning program is completed. In general, the summative test is carried out at the final test semester; therefore the summative test can be called the UAS. In other to achieve the targeted results, the summative test can be filled with tasks that have a low level of difficulty that have been tasted in the formative test.

Summative test is carried out after the end of giving a group or a larger program. The main purpose of the summative test is to determine the value that symbolizes the success of students after they have taken the learning process within a certain period of time.

The benefits of the summative tests are as follows:

- i. Determine the position or rank of each student in the group.
- ii. Determine whether or not students can continue the next learning program.
- iii. Inform students' progress to be conveyed to other parties such as parents and schools. If the summative test is carried out at the end of each semester, then at the end of each level of education a final test is carried out or commonly

¹⁵ H.Douglas Brown, *Language Assessment Principle and Classroom Practices* (New York: Pearson Education Inc, 2004).

called the final stage of learning evaluation (to fill in the record of student learning progress that student has achieved).¹⁶

Anas Sudijono revealed a formative test of learning outcomes which aims to determine the extent to which students "have been formed" (according to the learning objectives within a certain period of time).¹⁷ Meanwhile, Ngalm Purwanto which feedback, seek feedback which then the results can be used to improve the learning process that is being or as been done.

Formative test is usually carried out in the middle of learning, which is carried out every time the material or sub-topic end. Follow-up that can be done after knowing the results of the formative test if the results show that the material has been mastered well then it can be continued on the next subject but if the material has not been mastered well then the parts that have not been mastered can be developed or explained again.¹⁸

3. Characteristics of a good test

The characteristics of a good test as an evaluation tool are to meet the requirements of validity and reliability.¹⁹ Test is said to have good quality when have validity, reliability, high discrimination index, medium difficulty, and a working distracter.²⁰

¹⁶ Riinawati, *Pengantar Evaluasi Pendidikan* (Banjarmasin: Thema Publishing, 2021).

¹⁷ Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: PT Raja Grafindo Persada, 2012).

¹⁸ ErwinTri Wahyuningsih, "Analisis Butir Soal Tes Objektif Buatan Guru Ulangan Semester Ganjil Mata Pelajaran Ekonomi Kelas" (Universitas Negeri Yogyakarta, 2015).

¹⁹ Ina Magdalena et al., "Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan" 3 (2021): 198–214.

²⁰ Anak Agung Sri Dwipayani, "Analisis Validitas Dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Bahasa Indonesia Kelas X.D SMAN 1 Terhadap Pencapaian Kompetensi," 2013, 1–18.

a. Validity

1) Definition of validity

Validity is a measuring instrument that shows the level of a measuring instrument. A measuring instrument is less valid if it has low validity. To test the validity of a measuring instrument, the correlation price between the parts of the measuring instrument as a whole is sought by correlating each item of the measuring instrument with the total score, which is the sum of each item score.²¹

2) Factors that affect the validity of the test results

a) Evaluation instrument

Developing an instrument is not easy, especially if a teacher does not understand the procedures and evaluation techniques themselves. If the evaluation instrument is not good, it can result in poor evaluation results. For this reason, in developing an evaluation instrument, a teacher must pay attention to matters that affect the validity of the instrument and are related to the procedure for preparing the instrument, such as the syllabus, question grids, instructions for working on questions, filling out answer sheets, answer keys, use of effective sentences, alternative forms of answers, etc.

b) Evaluation and scoring administration

In the administration of evaluation and scoring, there are many irregularities or mistakes, such as disproportionate allocation of time for

²¹ Sri Sujarwadi, "Validitas Dan Reliabilitas Instrumen Penelitian," 2011, 1–23.

working on questions, providing assistance to students in various ways, students cheating on each other during exams, scoring errors, and the physical and psychological conditions of students who are less profitable.

c) Factor answers from students

The student answer factor is more influential than the previous two factors. This factor includes the tendency of students to answer quickly but not precisely, the desire to do trial and error, and the use of certain language styles in answering essay questions.²²

3) Valid test criteria

Validity involves the degree to which the measuring instrument's accuracy and precision enable it to perform its purpose. when a test or measurement tool delivers result that are accurate or consistent with the intended use of the measurement, this is referred to as have high validity This means that the measurement results from these measurements are the exact facts or actual condition of what is being measured. The question is declared invalid if the result of the calculation is < 0.3 and the question is declared valid if the result is ≥ 0.3 .

b. Reliability

1) Definition of reliability

Reliability is defined by the consistency of tests that are tasted repeatedly whose results are relatively the same, meaning that after the results of the first test with the next test are correlated, there is a significant

²² Muhammad Ropii and Muh Fahrurrozi, *Evaluasi Hasil Belajar*, ed. Syukrul Hamdi (Lombok: Universitas Hamzanwadi Press, 2017).

correlation result. Reliability refers to the tool in assessing what is desired, meaning the ability of the tool to be used to provide relatively the same result. Reliability is the consistency of the measurement of measuring instruments. This can be a measurement from the same measuring instrument that gives the same result or a more subjective measurement of whether two raters give similar scores.

2) Reliable test criteria

Test instrument can be trusted if it gives consistent results when tested many times. If students are given the same test at different times, then each student will remain in the same order. Reliability means how consistent the test scores are from one measurement to the next.

The degree of this relationship is indicated by the reliability coefficient that moves from 0 to 1. If the coefficient is getting closer to 1, then it is reliable and vice versa. It generally provides a minimum standard of reliability coefficient equal to or greater than 0.6. The question is declared unreliable if the result of the calculation is < 0.70 and the question is declared reliable if the result is ≥ 0.70 .

c. Level of difficulty

1) Definition of level of difficulty

According to Lyle F. Bachman, item difficulty is defined as the proportion of test takers who answered the item correctly, and the item difficulty index, values can be calculated on the basis of test takers'

response to the item.²³ The level of difficulty of an item is the proportion between the number of test takers who answered the item and the number of test takers. This means that if many test takers answer the item correctly, the greater the index of difficulty level, which means the easier the item will be. On the other hand, the fewer test takers who answer the item correctly, the more difficult the question will be.

2) Good test criteria based on level of difficulty

A good item is a question that is not too easy or not too difficult.²⁴ Problems that are too easy do not stimulate students to increase their efforts to solve them. On the other hand, questions that are too difficult will cause students to become desperate and not have the enthusiasm to try again because they are out of reach.²⁵ The difficulty level of the items is measured by the percentage of students who answer the questions correctly. If the questions are easy, the difficulty index is higher.

The level of difficulty of an item test can be caused by the complexity of the subject matter and the nature of the answer choices provided. The questions often confuse the test takers and the alternative answers are homogeneous or the question sentences are difficult to understand or have multiple meanings for the test takers.

²³ Lyle Bachman, "Statistical Analyses for Language Assessment (Cambridge Language Assessment)," 2004.

²⁴ Suharsimi, Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, 2nd ed. (Jakarta: Bumi Aksara, 2013).

²⁵ Abet Yani, Ali Fikri Asri, and Ahmad Burhan, "Analisis Tingkat Kesukaran, Daya Pembeda Dan Fungsi Distraktor Soal Ujian Semester Ganjil Mata Pelajaran Produktif Di SMK Negeri 1 Indralaya Utara," 2013, 98–115.

Based on the theories that have been put forward, the level of difficulty of the questions is how easy or difficult a question is for students. The level of difficulty is expressed by the percentage of students who answer correctly. The greater the percentage of students who answered the question correctly, the easier it was. On the other hand, the smaller the percentage of students who answered the question correctly, the more difficult it was.

The difficulty level criteria that used are as follows:

- a) An item with a difficulty level of 0.00 - 0.30 is a difficulty item
 - b) An item with a difficulty level of 0.31- 0.70 is a moderate item
 - c) An item with a difficulty level of 0.71 – 1.00 is an easy item
- d. Discrimination index
- 1) Definition of discrimination index

According to Daryanto, discrimination index is the ability of each item to distinguish between students who have low abilities and high abilities. Arifin added that the calculation of item discrimination is to measure the extent to which an item is able to distinguish students who have mastered competence from students who have not or less mastered competence based on certain criteria.²⁶ This opinion can be interpreted if it is given to students who are able to show high achievement and if given to students who are weak in knowledge the results are low.

²⁶ Arifin Zainal, *Evaluasi Pembelajaran*. (Bandung: PT Remaja Rosdakarya, 2014).

2) Purpose of discrimination index

Discriminatory analysis aims to determine the ability of the questions in distinguishing students who are classified as capable (high in achievement) with students who are classified as low in achievement. A test item has a discrimination index if the test item can be answered correctly by lower group students and cannot be answered correctly by upper group students. The higher the discrimination index of the question, the better the question in question distinguishes students who already understand and do not understand the material.

3) Good test based on discrimination index

The discrimination index of an item can be known by looking at the size of the item discrimination index number, which is a number that shows the size of the discrimination index of an item. If items can be answered correctly by smart students and not by smart students, then the item is not good because it has no discrimination index. Likewise, if all students, smart or not, cannot answer correctly, the item is not good, also because it does not have a discrimination index. Good items are items that are answered correctly by smart students. All students were grouped into two groups: the proficient group, or upper group, and the unskilled group, or lower group.

If the entire upper group can answer the question correctly while the lower group answers incorrectly, then this item has the largest discrimination index, which is 1.00. On the other hand, if all the upper groups answer incorrectly but all the lower groups answer correctly, then the value of the discrimination index is -1.00. But if students in the upper

group and students in the lower group both answered correctly or both answered incorrectly, then the question has a discrimination index value of 0.00 because it has no discrimination index.²⁷

e. Distractors

The distractors are important components of an item, as they show a relationship between the total test score and the distractor chosen by the student.²⁸ Student's performance depends upon how the distractors are designed. Distractor efficiency is one such tool that tells whether the item was well constructed or failed to perform its purpose.

The distractor has a number of options ranging from three to five, and of the possible answers attached to each item; one of them is the correct answer while the rest are wrong answers. Those wrong answers are known as distractors. The main purpose of installing a distractor for each item is so that the many students who take the learning achievement test are interested in choosing it because they think the distractor they have chosen is the right answer.²⁹

According to Mozaffer, the design of the distractions affects a student's performance. This is why it's crucial to consider the Distractor Efficiency, which shows whether or not the item's distractions were successful in keeping students from choosing the correct response. Every distraction must be relatively close to the item's key. Functional

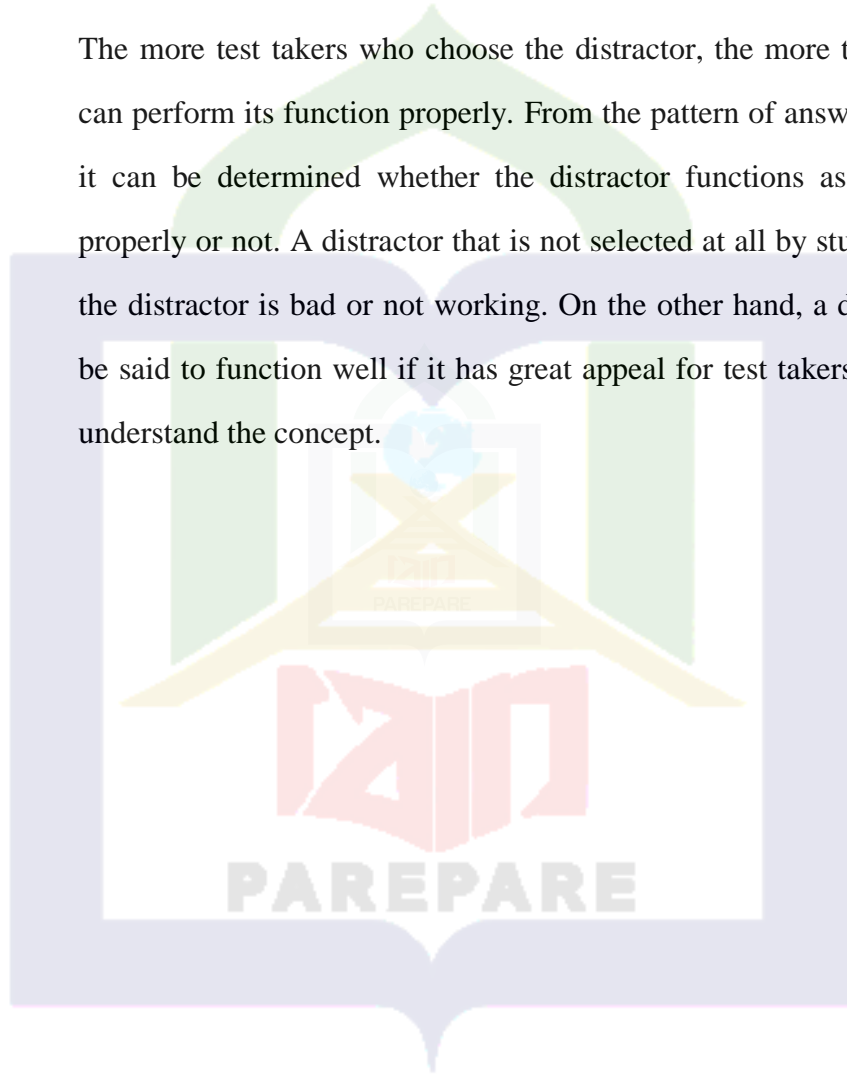
²⁷ Joko Widiyanto, *Evaluasi Pembelajaran* (Madiun: Unipma Press, 2018).

²⁸ Ismail Burud, Kavitha Nagandla, and Puneet Agarwal, "Impact of Distractors in Item Analysis of Multiple Choice Questions" 7, no. 4 (2019): 1136–39.

²⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: PT Raja Grafindo Persada, 2012).

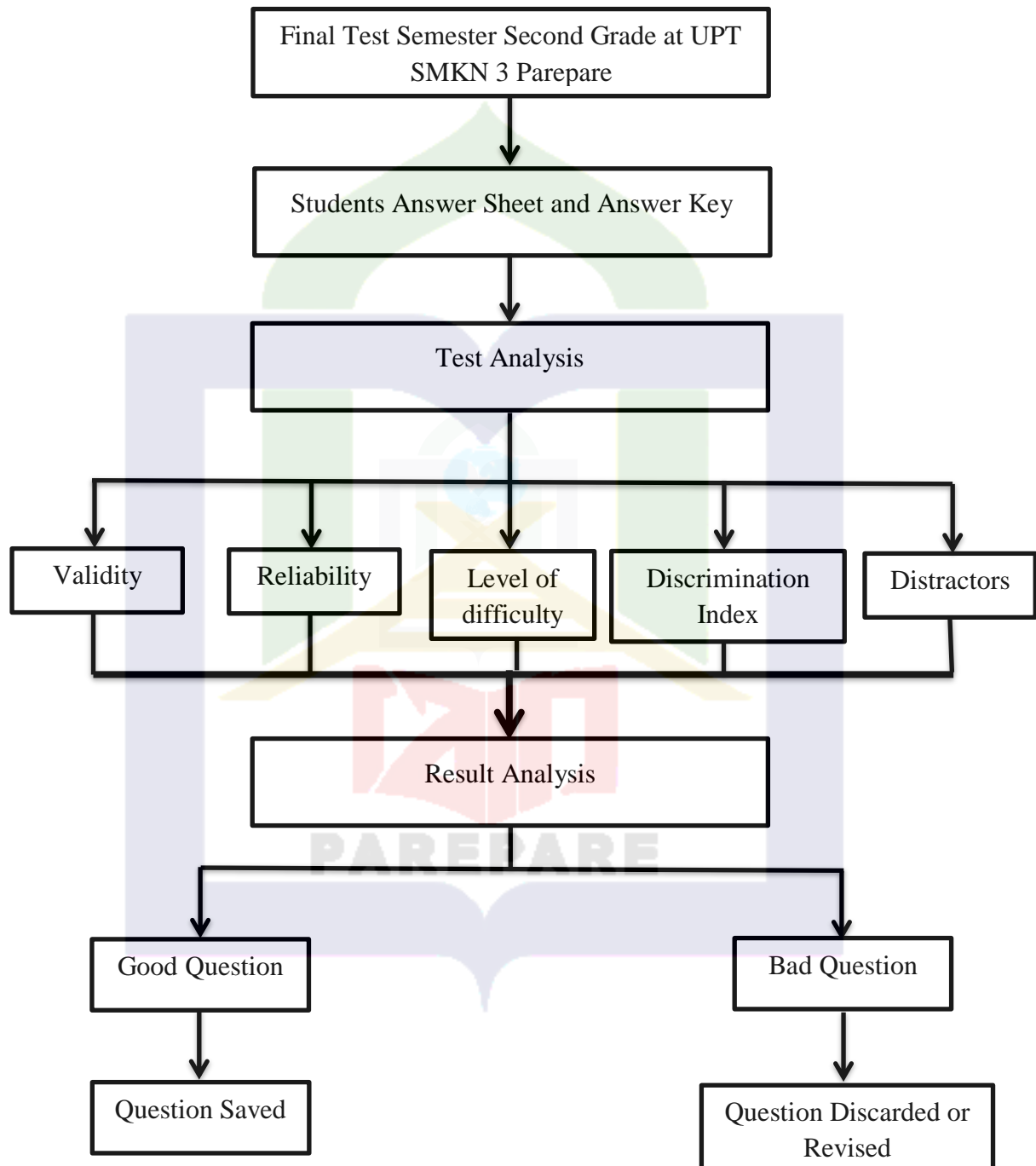
distractions are those that $>5\%$ or more of the test-takers choose, and nonfunctional distractions are those that 5% of the test-takers choose.³⁰

The effectiveness of the distractor is how well the wrong choice can deceive the test taker who does not know the available answer key. The more test takers who choose the distractor, the more the distractor can perform its function properly. From the pattern of answer questions, it can be determined whether the distractor functions as a distractor properly or not. A distractor that is not selected at all by students means the distractor is bad or not working. On the other hand, a distractor can be said to function well if it has great appeal for test takers who do not understand the concept.



³⁰ Mozaffer Rahim Hingorjo and Farhan Jaleel, "Analysis of One-Best MCQs : The Difficulty Index , Discrimination Index and Distractor Efficiency," *2012* 62, no. 2 (n.d.): 142–47.

C. Theoretical Framework



D. Hypothesis

Based on the theoretical basis that has been presented previously, the researcher can formulate the research hypothesis as follows.

1. The validity of the English teacher-made final test semester for second grade of UPT SMKN 3 Parepare for the academic year 2022/2023 was declared valid.
2. Reliability level the English teacher-made final test semester of UPT SMKN 3 Parepare for the academic year 2022/2023 has a high level of reliability.
3. Difficulty level of the English teacher-made final test semester for second grade of UPT SMKN 3 Parepare for the academic year 2022/2023 is in accordance with the proportion of the level of the final exam of difficulty in general.
4. The English teacher -made final test semester for second grade English at UPT SMKN 3 Parepare for the academic year 2022/2023 has good distinguishing power.
5. The effectiveness of the English teachers-made final test semester for second grade English test at UPT SMKN 3 Parepare for the academic year 2022/2023 can work well.

CHAPTER III

METHODOLOGY OF RESEARCH

A. The Research Design

The type of research used in this research was quantitative descriptive. Quantitative Descriptive is research aims to describe the situation precisely and accurately, not to look for a relationship between the independent variable and the dependent variable or to compare two or more variables to find causation.³¹ Quantitative descriptive is a research that describes and analyzes data in the form of an English teacher-made test. This research approach used a quantitative approach because the information was realized in the form of numbers.

B. Location and Time

1. Location

The research location was UPT SMK 3 Parepare in Jl. Karaeng Burane No.16, Ujung District, Parepare, and South Sulawesi. The researcher chose this location because one of the English teachers at this school never conducted an analysis of the test given to second grade students. For this research, the researcher chose second grade students as research subjects, the specific sample for this research was class XI Multimedia 1.

2. Time

The implementation of this research took place over a period of 1 month, from May to June 2023, to collect data.

³¹ Ratna Wijayanti Daniar Paramita, Noviansyah Rizal, and Riza Bahtiar Sulistyan, *Metode Penelitian Kuantitatif* (Widya Gama Press, 2021).

C. Population and Sample

1. Population

The population of this research was all students of class XI Multimedia at UPT SMKN 3 Parepare for the 2022/2023 academic year, consisting of 101 students. The researcher selected from a set of the final test, made by one of the English teacher in UPT SMKN 3 Parepare named St. Ruwaedah, because only test she made had multiple-choice, and the teacher did not know how to analyze the test based on validity, reliability, level of difficulty, discrimination index, and distractors. The set of final test chose by the researcher for this research was a set of summative test.

This final exam consists of 15 questions, 10 of which are multiple-choice and 5 of which are essay questions. In order to get the data, and the answers can be more easily corrected, the writer only analyze the test in the form of multiple-choice. To get the data in the form of test scores, the researcher carried out the result of the students' works. The students' works are the students' answers which recorded in the answer sheets.

2. Sample

In this study, the researcher used a purposive sample. Purposive sample is a sample selected based on certain considerations with the aim of obtaining a sample that has the desired characteristics.³² The total number of students in class XI multimedia was 101, but I not took the whole class as my sample. I selected XI Multimedia 1 with 33 students.

³² Nugraha Setiawan, "Teknik Sampling," 2005, 25–28.

D. Technique of Data Analysis

The data in this study was collected used documentation and the formula used in calculate the validity, reliability, level of difficulty, discrimination index, and distractors. The documentation technique is the collection of research data through a number of documents (information obtained) in the form of written documents or recorded documents.³³ The documentation used to collect data is in the form of final questions made by an English teacher, answer keys, and student answers so that they can be directly used as data analysis material to analyses the questions.

E. Instrument of the research

In this study, the technique used by researchers is documentation. Documentation is data collection by viewing data or information by studying written or recorded data. The documentation method used to obtain data is test made by an English teacher, answer key, and student answers.

F. Operational Definition of Variables

Item analysis is an assessment of each test question to obtain good quality questions in order to measure students' abilities. Aspects that need to be considered in carrying out item analysis are as follows:

1. Validity

Item validity is the ability to measure exactly test made English teacher. A test or measurement device is said to have high validity if the test

³³ Rahmadi, *Pengantar Metodologi Penelitian*, ed. Syahrani (Banjarmasin: Antasari Press, 2011).

performs the measuring function correctly or provides measurement results that are in accordance with the purpose of the measurement.

2. Reliability

An indicator of how much a measure device can be trusted or relied upon is called reliability. If a test yields the same findings when administered to same group of students over different period, it is said to be reliable.

3. Level of Difficulty

Level of Difficulty is a measure that indicates how difficult or easy a test is. If many test takers answer the item correctly, the greater the index of difficulty level, which means the easier the item will be. On the other hand, the fewer test takers who answer the item correctly, the more difficult the question will be. Good questions are questions that are not too easy and not too difficult.

4. Discrimination index

Discrimination index is the ability of a question to distinguish between students who are able and students who are less able to answer the question. The discriminatory power of questions will examine test in terms of the test's ability to distinguish students who are in the high and low achievement categories.

5. Distractors

The effectiveness of the distractor is the distribution in determining the answer choices on multiple-choice questions. The pattern of the answer to the question can determine whether the distractor is functioning properly or not.

G. Technique of Data Analysis

Analysis of test items to obtain important information for teachers about the quality that has been made. Item analysis is carried out by examining the validity, reliability, level of difficulty, discrimination index, and effectiveness of distractors.

1. Validity

To calculate the validity of items on multiple choice questions, the researcher used the following formula:

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Where:

- r_{pbi} = Biserial point correlation
- M_p = Mean of students who answered correctly
- M_t = Total mean score
- S_t = Total standard deviation
- p = Proportion of correct answers to all answers
- q = Proportion of incorrect answers to all answers
- $r_{pbis} < 0.3$ = invalid
- $r_{pbis} \geq 0.3$ = valid³⁴

2. Reliability

To calculate the reliability of items on multiple choice questions, the researcher used the following formula:

$$KR - 20 = \frac{n}{n - 1} \left(\frac{S_t^2 - \sum pq}{S_t^2} \right)$$

³⁴ Abdul Qodir, *Evaluasi Dan Penilaian Pembelajaran*, ed. Ngalimun, Yogyakarta (K-Media, 2017).

Where:

KR = Kuder Richardson

n = number of items in the test

S_t^2 = the variance of the total test score of the square of standard deviation

p = Proportion of correct answers to all answers

q = Proportion of incorrect answers to all answers

KR-20 \geq 0.70 = reliable

KR-20 $<$ 0.70 = unreliable³⁵

3. Level of difficulty

To calculate the level of difficulty of items on multiple-choice questions, the researcher used the following formula:

$$P = \frac{B}{JS}$$

Where:

P = level of difficulty

B = the number of students who answer the items correctly

JS = number of students in each group

The difficulty level criteria that used are as follows:

1. An item with a difficulty level of 0.00 - 0.30 is a difficulty item
2. An item with a difficulty level of 0.31- 0.70 is a moderate item
3. An item with a difficulty level of 0.71 – 1.00 is an easy item³⁶

³⁵ Fajri Ismail, Mardiah Astuti, and Hani Atus Sholikhah, *Evaluasi Pembelajaran Berbasis Riset* (Palembang: Karya Sukses Mandiri, 2020).

4. Discrimination index

To calculate the discrimination index of items on multiple choice questions, the researcher used the following formula:

$$DP = \frac{WL - WH}{n}$$

Where:

DP = Discriminating Power

WL = the number of students in the lower group who answer the items incorrectly

WH = the number of students in the upper group who answer the items incorrectly

n = 27 % X N

Table 3.1. Criteria Used in Analyzing Discrimination Index

DP	Quality	Recommendations
0,70 -1,00	Excellent	Retain
0,40 – 0,69	Good	Possible to improve
0,20 – 0,39	Satisfactory	Need to check/review
< 0,20	Poor	Discard/review in depth
<0,00 (-)	Worst	Definitely discard ³⁷

5. Distractors

Answer option is distribution of students in terms of determining answer choices by counting the number of students who choose answer choices a,b,c,d or who do not choose any option called omit. The answer

³⁶ Arief Aulia Rahman and Cut Eva Nasryah, *Evaluasi Pembelajaran* (Uwais Inspirasi Indonesia, 2019).

³⁷ Muhammad Ropii and Muh Fahrurrozi, *Evaluasi Hasil Belajar*, ed. Syukrul Hamdi (Lombok: Universitas Hamzanwadi Press, 2017).

options for questions can be determined whether the distractor is functioning properly or not. Distractor that is not chosen at all by students means that the distractor does not work. On the other hand, a distractor can be said to function properly if it has a great appeal for students who do not understand the concept and do not understand the material.

The following criteria can be used to calculate the effectiveness of the question distractor:

- i. The effectiveness of the distractors is said to be very good if the four distractors work.
- ii. The effectiveness of the distractor is said to be good if there are three distractions that function.
- iii. The effectiveness of the distractor is said to be quite good if there are two distractions that function.
- iv. The effectiveness of the distractor is said to be poor if there is one distractor that functions. The effectiveness of the distractor is said to be bad if all four (all) distractors don't work.³⁸

³⁸ Sugiyono, *Metode Penelitian Pendidikan Pendekatan Kuantitatif Dan R&D* (Bandung: Alfa Beta, 2010).

CHAPTER IV

FINDINGS AND DISCUSSIONS

A. Findings

Item analysis is the study of test questions in order to obtain a set of questions that have adequate quality.³⁹ In addition, item analysis was carried out to find errors or mistakes in making the test. This research is based on the analysis of the final multiple-choice test for the XI Multimedia Class English subject at UPT SMKN 3 Parepare, South Sulawesi. In this study, item analysis was carried out based on validity, reliability, level of difficulty, discrimination index, and distractors. The number of questions is 10, and there are 33 students in class XI Multimedia 1.

In this study, the instrument used was documentation to obtain data in the form of a final test, answer keys, and student answer sheets. The way to analyse the problem was to use a formula that has been determined by experts.

The following are the results of research conducted at SMKN 3 Parepare Class XI Multimedia 1. The scores obtained from the results of student answers to a total of 10 multiple-choice questions had a correct answer weight of 1 and an incorrect answer weight of 0.

³⁹ Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar* (Bandung: Remaja Rosdakarya, 2014).

Table 4.1. Student Response Score XI Multimedia 1 on the Final Test Multiple-Choice at UPT SMKN 3 Parepare

NO.	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
1	Elif Hartiningsih	1	1	1	1	1	1	1	0	1	1	9
2	Aisyah Fadhillah Muh. Azis	1	1	1	1	1	1	1	0	1	1	9
3	Astri Cindy Manchetti	1	1	1	1	1	1	1	0	1	1	9
4	Andi Mahathir Muhammad	1	1	1	1	1	1	1	0	1	0	8
5	Ayu Dwi Lestari	1	1	1	0	1	1	1	0	1	1	8
6	Amelya Nanda	1	1	1	1	1	0	1	0	1	1	8
7	Aisyah Atirah	1	1	1	1	1	1	0	0	1	1	8
8	Chelsea Olivia Harum	1	1	0	1	1	1	1	0	1	1	8
9	Dirham Dwi Yulianto Fereira	1	1	0	0	1	1	1	0	1	1	7
10	Fito Anugrah Nurzaman	1	1	1	1	1	0	0	0	1	1	7
11	Airin Tenriani	1	1	0	0	1	1	1	0	1	1	7
12	Bania Nur Sakinah	1	1	0	1	0	1	1	0	1	1	7
13	Akhiryati Berlian	1	0	1	1	1	1	0	1	1	0	7
14	Audya Pertiwi	1	0	1	1	1	0	1	1	0	1	7
15	Dian Sardi Jay	1	1	0	0	1	1	1	1	1	0	7
16	Azyurah Dwi Putri	1	1	1	0	0	1	0	1	1	1	7
17	Alfiana Muhsin	1	0	1	1	1	0	1	0	0	1	6
18	Alya Sephira Fitriani	1	1	1	1	1	0	0	0	1	0	6
19	Citra Angreni	1	1	1	0	1	1	0	0	1	0	6
20	Acyuta Mujahid. S	1	1	1	1	1	1	0	0	0	0	6
21	Farah Quin Afriyani	1	0	1	0	1	0	1	1	0	1	6
22	Astri	1	0	0	1	1	1	0	0	1	0	5
23	Aulia Rahman	0	1	0	1	1	1	1	0	0	0	5
24	Deby Grace Mandri	0	0	1	0	1	1	1	1	0	0	5
25	Indah Suci Amelia	1	0	0	1	1	0	0	0	1	0	4
26	Aisyah Ramadhani	0	1	0	0	1	1	0	1	0	0	4
27	Emi	1	0	0	1	1	0	1	0	0	0	4
28	Ayu Andini	0	1	0	1	0	1	0	0	0	0	3
29	Arnita Melani	1	0	0	0	0	0	0	1	0	0	2
30	Adam Subarkah	0	1	0	0	1	0	0	0	0	0	2
31	Febriyani Nurul Aini Pipin	0	0	0	0	0	0	1	0	0	1	2

Continued Table 4.1

NO.	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
32	Andini Pertiwi	0	0	0	1	0	0	0	0	0	0	1
33	Delia Istiqinah	0	0	0	0	0	0	0	0	0	1	1
	TOTAL	25	21	17	20	26	20	18	8	19	17	191

1. Validity

In testing validity, researchers used the Biserial Point Correlation formula. This calculation used criteria that are below 0.30 when entering invalid criteria and above or equal to 0.30 when entering valid criteria. Based on the results of the analysis with 10 multiple-choice questions and 33 students in the XI Multimedia Class English subject at UPT SMKN 3 Parepare, the results of the validity test showed that 2 or 20% of the items are invalid and 8 or 80% of the items are valid.

Table 4.2 The result of an English Teacher-Made Test Analysis Based on the Validity.

No	Category	Item	Total	Percentage
1.	Valid $r_{pbis} \geq 0.3$	1, 2, 3, 5, 6, 7, 9, and 10.	8	80%
2.	Invalid $r_{pbis} < 0.3$	4 and 8	2	20%

So, items 1, 2, 3, 5, 6, 7, 9, and 10 can or are suitable for reuse, and these items must be maintained by saving them in the question bank because included in category good question based on validity. Items that are invalid should be corrected by increasing the teacher's technical mastery in preparing the items. The item can be

valid because the preparation of the questions is good and includes material that really represents the target measure.

2. Reliability

In calculating reliability, this study used the KR-20 formula. Testing the reliability of the questions based on the guidelines: if the value of the calculated reliability is greater than or equal to 0.70, then the questions tested have high reliability, but if the calculated reliability values are less than 0.70, then the questions tested have low reliability or are not reliable.

Based on the KR-20 formula for calculating reliability, the result of the multiple choice test analysis made by one of the teachers at UPT SMKN 3 Parepare is 0.6. Based on the results of these calculations, it can be concluded that the multiple-choice test is unreliable.

3. Level of difficulty

Based on the formula used in calculating the level of difficulty, there are criteria for calculating these items. These criteria are: questions with a difficulty level of 0.00–0.30 are difficult questions; questions with a difficulty level of 0.31–0.70 are moderate questions, and questions with a difficulty level of 0.71–1.00 are easy questions.

Based on the results of the multiple-choice test analysis made by an English teacher at UPT SMKN 3 Parepare, which was answered by students of class XI Multimedia 1, the results obtained from 33 students showed that 1 item was difficult to answer, 8 items were in the moderate category, and 1 item was in the easy category. This means that 8 items that can be reused.

Table 4.3 The result of an English Teacher-Made Test Analysis Based on the Level of Difficulty.

No	Category	Item	Total	Percentage
1.	Difficult (0.00 – 0.30)	8	1	10%
2.	Moderate (0.31 – 0.70)	1, 2, 3, 4, 6, 7, 9, and 10	8	80%
3.	Easy (0.71 – 1.00)	5	1	10%

Items that fall into the moderate category must be maintained and stored in the question bank because these items can or deserve to be reused and included in category good question based on level of difficulty. Items in the difficult category must be re-examined in order to identify the factors that make it difficult for students to answer them. Items that fall into the easy category should be corrected or researched so that the factors that make almost all students answer correctly can be identified. Improvements can be made by improving answer options or making sentences more complex. After correcting the items, they can be saved to the question bank.

4. Discrimination index

Based on the formula used to calculate the discrimination index, there are several criteria, including: if the discrimination index is 0.70–1.00, then it is included in the excellent category; 0.40–0.69 is good, 0.20–0.39 is satisfactory, < 0.20 is poor, and < 0.00 (-) is the worst category.

Based on the results of the analysis on the multiple-choice test made by the English teacher at UPT SMKN 3 Parepare, which was answered by students of class XI Multimedia 1, there were 2 questions that were included in the excellent category, 6 questions that were in the good category, 1 item that was included in the satisfactory category, and 1 item that was included in the worst category. This means 2 items that cannot be reused.

Table 4.4 The result of an English Teacher-Made Test Analysis Based on Discrimination Index

No	Category	Item	Total	Percentage
1.	Excellent (0.70 – 1.00)	3 and 9	2	20%
2.	Good (0.40 – 0.69)	1, 2, 5, 6, 7 and 10	6	60%
3.	Satisfactory (0.20 – 0.39)	4	1	10%
4.	Worst (< 0.00 / -)	8	1	10%

Items with excellent and good discrimination index must be maintained by including them in the question bank because these items can or deserve to be reused and included in category good question based on discrimination index, while items with sufficient and worst discrimination index must be totally repaired by tracing the cause of the failure. One effort that can be made is to fix questions that are unclear in their meaning so that they confuse students who are highly capable of answering them. Items must be able to provide differences in ability between students who understand the material and students who do not understand the material.

5. Distractors

There are criteria to determine whether the distractor is working or not. These criteria are: the effectiveness of the distractors is said to be very good if the four distractors work, the effectiveness of the distractor is said to be good if there are three distractions that function, the effectiveness of the distractor is said to be quite good if there are two distractions that function, the effectiveness of the distractor is said to be poor if there is one distractor that functions, the effectiveness of the distractor is said to be bad if all four (all) distractors don't work.

Table 4.5 The Number of Options Selected by Students

Item	Option					Total
	A	B	C	D	E	
1	3	2	*25	3	0	33
2	1	6	3	*21	2	33
3	3	10	1	2	*17	33
4	*20	1	1	11	0	33
5	3	*26	1	2	1	33
6	7	2	*20	2	2	33
7	13	*18	0	2	0	33
8	8	*8	11	4	2	33
9	0	5	3	*19	6	33
10	4	4	6	2	*17	33

In a multiple-choice test made by an English teacher at UPT SMKN 3 Parepare and answered by students XI Multimedia 1, 7 distractors function very well because the 4 distractors were chosen by students who answered the test. Meanwhile, 2 items have good distractors because 1 distractor does not work or not chosen by students. Besides that, there is also 1 item that is quite a good distractor because there

are 2 distractors that don't work. To conclude, only 1 item that cannot be reused because has category quite a good distractor.

Table 4.6 The Result of an English Teacher-Made Tests Analysis Based on Distractors

No	Category	Item	Total	Percentage
1.	Very good	1, 2, 3, 5, 6, 8 and 10	7	70%
2.	Good	4 and 9	2	20%
3.	Quite good	7	1	10%

Items whose distractors function very good and good must be maintained are put in the question bank because these items can or deserve to be reused and included in category good question based on distractor. Items whose distractors function quite good should be corrected so that these distractors become good questions. Corrective steps can be taken by replacing the distraction.

B. Discussion

Based on the results of the analysis of multiple choice-items made by an English teacher at UPT SMKN 3 Parepare and answered by students of class XI Multimedia 1, this discussion discussed in detail the results of the analysis in terms of validity, reliability, level of difficulty, discrimination index, and effectiveness of the distractor.

1. Validity

A test can be said to be valid if it can measure the object that should be measured in accordance with the criteria. Biserial Point Correlation index obtained from the results of calculations used the criteria, which are: 0.30 is considered invalid

and above or equal to 0.30 is considered valid. So this validity is calculated using the point biserial correlation formula through the Biserial Point Correlation.

Based on the results analysis validity of the multiple-choice items for English subjects made by an English teacher at UPT SMKN 3 Parepare and which have been answered by 33 XI Multimedia 1 students, it shows that the multiple-choice questions that are included as valid questions are 8 or 80% items, and 2 or 20% of the items are invalid. This means that most of the items are valid. The results of this study were almost the same as the results of research from Lukmanul Hakim and Irhamsyah which showed that there were 23 valid items (92%) and 2 items (8%) included in the invalid category.⁴⁰

Valid items mean that the item has been able to carry out its function, namely being able to measure what should be measured. This is in accordance with Anas Sudijono's statement, which argues that the validity of an item from a test is determined by the accuracy of the measure possessed in measuring what should be measured through the item.⁴¹ Then, invalid questions can be caused by various factors. This relates to the statement by Gronlund in Ropii and Farurrozi, which states that there are 3 factors that affect the validity of test results: the instrument factor used for the test, administration and scoring factors, and factors from student answers.⁴² In the final multiple-choice test at UPT SMKN 3 Parepare, which was answered by students of class XI Multimedia 1, it could have been influenced by the tendency of students to answer quickly and inaccurately.

⁴⁰Lukmanul Hakim and Irhamsyah, "The Analysis Teacher-Made Test for Senior High School at State Senior High School 1 Kutacane" 21, no. 1 (2020): 10–20.

⁴¹Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: PT Raja Grafindo Persada, 2012).

⁴²Muhammad Ropii and Muh Fahrurrozi, *Evaluasi Hasil Belajar*, ed. Syukrul Hamdi (Lombok: Universitas Hamzanwadi Press, 2017).

The items can be said to be valid because they are of good quality and include material that represents the target measure. Then, invalid items should be corrected immediately before being used again by increasing mastery of the item preparation techniques. In this case, the teacher should use a good tool or program. In addition, the teacher can also consult experts to determine the validity of the questions that have been made before they are used as evaluation material. Thus, the questions to be tested are truly valid and can be said to be good questions to be used as an evaluation tool.

2. Reliability

A good test must have a good degree of reliability. A test can be said to have a high level of reliability if it can provide consistent results. A test is said to have a good level of reliability if it is able to give consistent results even though it is tested on same subjects and different times. An analysis of the items in terms of reliability is very important because, with this analysis, we will be able to find out how much consistency or determination is needed to measure the questions so that they can be trusted.

The results showed that the final test for English subject Class XI Multimedia 1 UPT SMKN 3 Parepare for the academic year 2022-2023 has a reliability index of 0.6. Based on this, it can be said that the test is included in the unreliable category because the value is < 0.70 . This is related to the statement of Fajri Ismail and friends that If the KR-20 is equal to or greater than 0.70, it means that the learning outcomes test that is being tested for reliability has been declared to have high reliability or is reliable. Meanwhile, if it is smaller than 0.70, it means that the learning outcomes test

being tested for reliability does not yet have high reliability or is unreliable.⁴³ The results of this study are the same as the research conducted by Taufiq Effendi and Illza Mayuni which showed reliability results with a value of 0.48 which means unreliable.⁴⁴

Therefore, if this question is tested again in the same group, the results will not change. Based on this analysis, it can be interpreted that the test reliability value for the final test for English subject Class XI Multimedia 1 UPT SMKN 3 Parepare was still in the low category. This is probably because the test scores of students in this final test have a low level of diversity. As revealed by Brown there are the factors affecting to the reliability; first, the most common reliability is caused by test takers temporary illness, fatigue, a bad day, anxiety and other physical or psychological factors. Second, it may be as the result of human error, subjectivity and bias. The next is cause by the condition which the test is administered. And the last, the nature of test itself can because measurement errors such as the length of test, the ambiguity options etc.⁴⁵

3. Level of difficulty

The level of difficulty of an item is one way that can be used to find out how good the quality of the item is. If the difficulty level value is close to 0, then the item is too difficult, whereas if the difficulty level value is close to 1, then the item is too easy, so it needs to be discarded. The results showed that the final test for English

⁴³ Fajri Ismail, Mardiah Astuti, and Hani Atus Sholikhah, *Evaluasi Pembelajaran Berbasis Riset* (Palembang: Karya Sukses Mandiri, 2020).

⁴⁴ Taufiq Effendi and Ilza Mayuni, "Examining Teacher-Made English Test in a Language School" 2, no. 2 (2022): 67–76.

⁴⁵ H. Douglas Brown, *Language Assessment Principle and Classroom Practices* (New York: Pearson Education Inc, 2004).

subject Class XI Multimedia 1 UPT SMKN 3 Parepare for the academic year 2022-2023 consisted of 10% or 1 item included in difficult questions, 10% or 10 items included in easy questions, and 80% or 8 items included in the category of moderate questions.

Based on the description above, it can be concluded that the final test English subject Class XI Multimedia 1 UPT SMKN 3 Parepare for the 2022-2023 academic year includes questions that are of moderate quality based on their level of difficulty because 80% of the questions were included in the moderate criteria. The results of this study are the same as those conducted by Mutiara Kusumawati and Samsul Hadi, whose results of the analysis of the level of difficulty are more dominant in the medium category with a score of 60%.⁴⁶ The test that is used for final test purposes was supposed to use questions with a moderate level of difficulty. Suharsimi Arikunto stated that good questions are questions that are not too easy or too difficult. Questions that are too easy will not stimulate students to think or enhance their efforts to solve problems in each item. On the other hand, questions that are too difficult will cause students to become discouraged and have no enthusiasm to try again because they are out of reach.⁴⁷

Items that fall into the difficult or easy category will not be able to distinguish the ability of one student from that of other students. Therefore, good questions are questions that fall into the moderate category. There are several factors that cause items to have a difficult category, including : students do not understand the material

⁴⁶ Mutiara Kusumawati and Samsul Hadi, "An Analysis of Multiple Choice Questions (MCQs): Item and Test Statistics from Mathematics Assessments in Senior High School" 4, no. 1 (2018): 70–78.

⁴⁷ Suharsimi, Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, 2nd ed. (Jakarta: Bumi Aksara, 2013).

that has been taught by the teacher, item sentence structures that are difficult for student to understand and item distractors that are almost the same so student find it difficult to distinguish between right and wrong answers.

Based on the results of the analysis of the questions based on the level of difficulty, the follow-up that needs to be done is the first. For items based on the results of the analysis included in the good category, these items should be recorded immediately in the question bank book. Furthermore, these items can be issued again in tests of learning outcomes in the future.

Second, for items that fall into the too difficult category, there are three possibilities for follow-up, namely: (1) the item is discarded and will not be issued again in future learning outcomes tests; (2) re-examined, tracked, and traced so that the factors that cause the items in question to be difficult to answer by students can be known; Are the sentences in the questions unclear, are the instructions for how to work on or answer the questions difficult to understand, or are there unclear terms in the questions. After repairs are made again, these items will be issued again in the upcoming learning outcomes test. (3) It must be understood that not every item included in the too difficult category is totally useless. Items that are too difficult can still be used in tests (especially selection tests), which are very strict in nature.

Third, for items that fall into the too easy category, there are three possible follow-ups, namely: (1) the items are discarded and will not be issued again in future learning outcomes tests. (2) Re-examined, tracked, and traced carefully in order to find out the factors that caused these items to be answered correctly by almost all students; there is a possibility that the distractor attached to the item in question is too easily known by students. The teacher must try to improve or replace it with another

option in such a way that it is difficult for students to distinguish between the answer key and the distractor. After making improvements, the item in question is tried again on the next learning achievement test. (3) Items that are too easy also have benefits, namely that items included in this category can be used in tests that are loose in nature, in the sense that most students will pass the selection test. In conditions like this, it is very wise if the items issued in the selection test are items that are included in the too easy category, so that the selection test can be said to be just a formality.

4. Discrimination Index

Analyzing the items in terms of discriminating index was carried out with the aim of knowing the ability of the items to be tested in order to be able to distinguish between students who have not mastered or lack mastery of the material to be tested and students who have mastered the material to be tested. A positive discrimination index value means having a high discrimination index, while a negative discrimination index means having a low discrimination index.

In this analysis, those included in the positive discrimination index were in the good and sufficient categories, while those included in the negative discrimination index were items that were in the bad and very bad categories. Items that are supported by a good potential discrimination index will be able to distinguish students who have high abilities (intelligent) from students who have low abilities (less intelligent).

Based on the results of the analysis, it can be seen that the final test for English subject Class XI Multimedia 1 UPT SMKN 3 Parepare for the academic year 2022/2023 consists of 10% or 1 item in the worst category, 1 item in the category of

satisfactory, and good categories as much as 6 or 60% of items, and items that fall into the category of very good as much as 20% or 2 items. The good category is more dominant than the other categories. The results of this study are different from the research conducted by Maya Marsevani with the results of the excellent category being more dominant with a score of 80%.⁴⁸

Good discrimination index indicates that the questions are able to distinguish students who have the ability to answer questions from students who do not have the ability to answer questions correctly. This relates with statement by Joko Widiyanto, which stated that the higher the coefficient of discrimination index of an item, the more capable the item is of distinguishing between students who master competence and students who lack competence.⁴⁹ Meanwhile, discrimination index that was not good indicated that the questions could not differentiate between students who had the ability to answer questions and students who could answer questions correctly.

Based on the results of this analysis, it proves that most of the questions fall into the category of good discrimination index. This means that the questions are able to distinguish between students who are good at it and those who are not very good at it. Questions with negative discrimination index should not be used or just thrown away. Item that have the worst and satisfactory category because more students answered correctly in lower group than in upper group's answers mean that students answered the questions correctly just by guessing.

⁴⁸ Maya Marsevani, "Item Analysis of Multiple-Choice Questions : An Assessment of Young Learners" 10, no. 2 (2022): 401–8.

⁴⁹ Joko Widiyanto, *Evaluasi Pembelajaran* (Madiun: Unipma Press, 2018).

5. Distractors

An analysis of the items in terms of the effectiveness of the distractor was carried out with the aim of finding out how well the wrong choices of the answer questions can trick the test takers who do not know the available answer keys. There are five alternative answers, and of the five answers, there is one possible answer attached to each of these items, which is the correct answer (answer key), while the rest are wrong answers. Those wrong answers are called distractors. To find out the effectiveness of the distractor, look at the answer sheets for the students' questions. Based on the answer sheet, it will be known whether the distractor function can function properly or not.

Then related to the function of the distractor, namely, the distractor can function on each item when the results of each choice option on each item indicate very good or good. Meanwhile, the distractor is declared not working if the results of each option on each item state that it is quite good and poor. The item is said to be very good if the distractor for each item can function as a whole, namely all the answer options in each item. Then the items are declared good if the distractor for each item does not function as much as one alternative answer. Then each item is declared quite good if the distractor on each item does not function for as many as 2 alternative answers. Furthermore, the items are said to be poor if the distractor for each item does not function for as many as 3 alternative answers or all of them.

The results analysis showed that the final test items for English subject Class XI Multimedia 1 UPT SMKN 3 Parepare for the academic year 2022/202 in the form of multiple choice questions showed that 7 or 70% items very good quality

distractors, good quality category amounted to 2 or 20% items, and questions of quite good category amounted to 1 item or 10%.

Based on all the explanations above, it can be concluded that the final test for English Class XI Multimedia 1 UPT SMKN 3 Parepare for the 2022–202 academic year, based on the effectiveness of the distractor, are included in the category of very good questions, namely 70%. So it needs to be maintained in the very good category for the next test. This is related to the statement according to Sudijono which states that distractor that has been obtained function very well can be used again on future tests, while the distractors which are not functioning properly should be repaired or replaced with another distractor.⁵⁰ Whereas for items in the poor or not good category, they must be corrected if the questions are still to be used by replacing distractors that have more potential to be almost the same as the answers so that in choosing answers, students can think carefully and thoroughly in determining answers.

⁵⁰ Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: PT Raja Grafindo Persada, 2012).

CHAPTER V

CLOSING

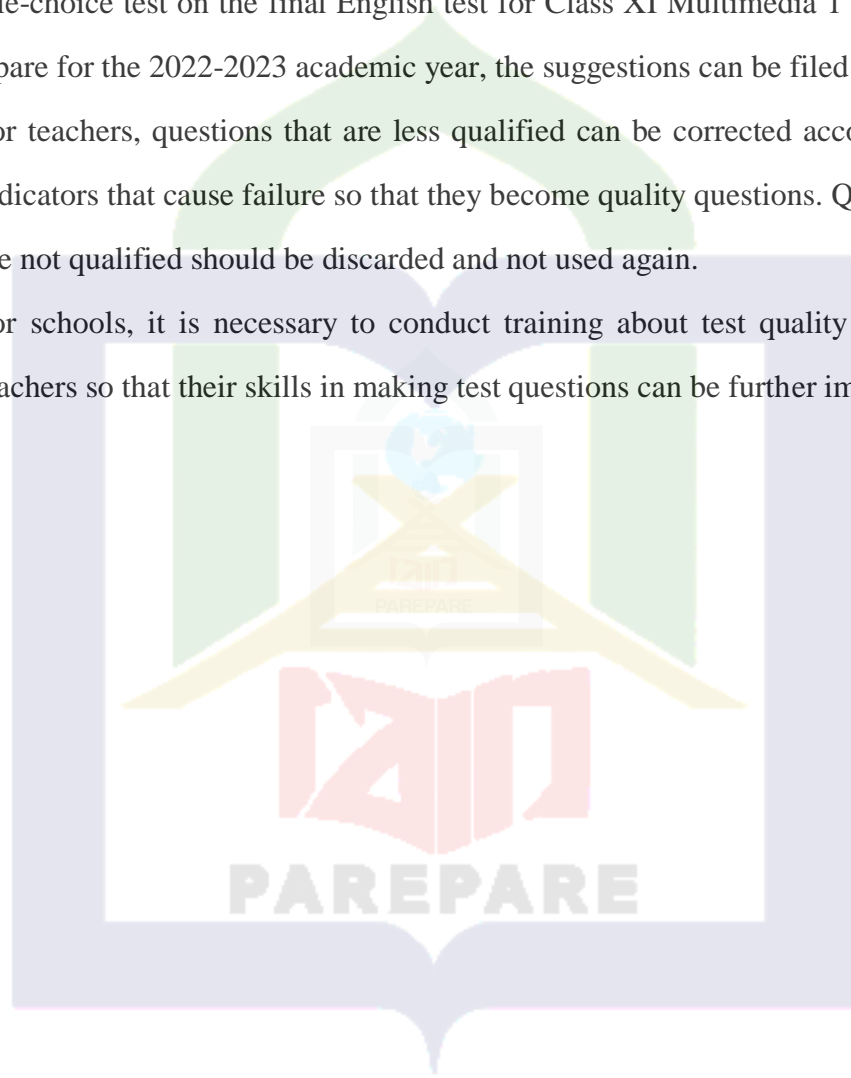
A. Conclusions

Based on the results of the research and discussion of item analysis, which includes terms of validity, reliability, level of difficulty, discriminating power, and effectiveness of the distractor in multiple-choice tests on the final English test for Class XI Multimedia 1 UPT SMKN 3 Parepare for the 2022-2023 academic year, it can be concluded that: Based on validity, the test included in the valid category is 8 items, while the questions that are invalid have 2 items. Based on reliability, included test that is unreliable with a value of 0.6. Based on the level of difficulty, the question included in the difficult category is 1 item, the moderate category is 8 items, and the easy category is 1 item. Based on the discrimination index, which includes questions that have the worst discrimination index is 1 item, satisfactory discrimination index category is 1 item, good discrimination index category is 6 items, and items that fall into the excellent discrimination index category are 2 items or 20%. Based on the effectiveness of the distractor, the items whose distractors function very good is 7 items, the distractors who are in the good category are 2 items %, 1 item is quite good category.

B. Suggestions

Based on the results of the item analysis consisted of validity, reliability, level of difficulty, discriminating power, and effectiveness of the contestants in the multiple-choice test on the final English test for Class XI Multimedia 1 UPT SMKN 3 Parepare for the 2022-2023 academic year, the suggestions can be filed as follows:

1. For teachers, questions that are less qualified can be corrected according to the indicators that cause failure so that they become quality questions. Questions that are not qualified should be discarded and not used again.
2. For schools, it is necessary to conduct training about test quality analysis for teachers so that their skills in making test questions can be further improved.



BIBLIOGRAPHY

Al-Qur'an Al-Karim

Adom, Dickson, Jephthar Adu Mensah, and Dennis Atsu Dake. "Test , Measurement , and Evaluation : Understanding and Use of the Concepts in Education" 9, no. 1 (2020): 109–19. <https://doi.org/10.11591/ijere.v9i1.20457>.

Asrul, Rusydi Ananda, and Rosnita. *Evaluasi Pembelajaran*. Medan: Citapustaka Media, 2014.

B, Nurgiyantoro. *Penilaian Pembelajaran Bahasa*. Yogyakarta: BPFE_Yogyakarta, 2010.

Bachman, Lyle. "Statistical Analyses for Language Assessment (Cambridge Language Assessment) (Lyle F. Bachman) (z-Lib.Org).Pdf," 2004.

Brown, H.Douglas. *Language Assessment Principle and Classroom Practices*. New York: Pearson Education Inc, 2004.

Burud, Ismail, Kavitha Nagandla, and Puneet Agarwal. "Impact of Distractors in Item Analysis of Multiple Choice Questions" 7, no. 4 (2019): 1136–39.

Dwipayani, Anak Agung Sri. "Analisis Validitas Dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Bahasa Indonesia Kelas X.D SMAN 1 Terhadap Pencapaian Kompetensi," 2013, 1–18.

Effendi, Taufiq, and Ilza Mayuni. "Examining Teacher-Made English Test in a Language School" 2, no. 2 (2022): 67–76.

Hakim, Lukmanul, and Irhamsyah. "The Analysis Teacher-Made Test for Senior High School at State Senior High School 1 Kutacane" 21, no. 1 (2020): 10–20.

Hartati, Neti, and Hendro Pratama Supra Yogi. "Item Analysis for a Better Quality Test," 2003, 12.

Haryanto, Kurniawan. "Analisis Butir Soal Pilihan Ganda Ulangan Akhir Semester Genap." Universitas Sanata Dharma, 2016.

Hingorjo, Mozaffer Rahim, and Farhan Jaleel. "Analysis of One-Best MCQs : The Difficulty Index , Discrimination Index and Distractor Efficiency." 2012 62, no. 2 (n.d.): 142–47.

Ismail, Fajri, Mardiah Astuti, and Hani Atus Sholikhah. *Evaluasi Pembelajaran Berbasis Riset*. Palembang: Karya Sukses Mandiri, 2020.

Kusumawati, Mutiara, and Samsul Hadi. "An Analysis of Multiple Choice Questions

- (MCQs): Item and Test Statistics from Mathematics Assessments in Senior High School” 4, no. 1 (2018): 70–78.
- Magdalena, Ina, Septy Nurul Fauziah, Siti Nur Faziah, and Fika Sulaehatun Nopus. “Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan” 3 (2021): 198–214.
- Marsevani, Maya. “Item Analysis of Multiple-Choice Questions : An Assessment of Young Learners” 10, no. 2 (2022): 401–8.
- Paramita, Ratna Wijayanti Daniar, Noviansyah Rizal, and Riza Bahtiar Sulistyan. *Metode Penelitian Kuantitatif*. Widya Gama Press, 2021.
- Putri, Yohana Fransisca Dewi Risna. “An Analysis of Teacher-Made English Final Second Semester Test for the Year Eleven Students Of SMAN 1 Ambarawa In The Academic Year Of 2008/2009 Based on the Representativeness of Content Standard.” Semarang State University, 2009.
- Qodir, Abdul. *Evaluasi Dan Penilaian Pembelajaran*. Edited by Ngalimun. Yogyakarta. K-Media, 2017.
- Rahmadi. *Pengantar Metodologi Penelitian*. Edited by Syahrani. Banjarmasin: Antasari Press, 2011.
- Rahman, Arief Aulia, and Cut Eva Nasryah. *Evaluasi Pembelajaran*. Uwais Inspirasi Indonesia, 2019.
- Riinawati. *Pengantar Evaluasi Pendidikan*. Banjarmasin: Thema Publishing, 2021.
- Ropii, Muhammad, and Muh Fahrurrozi. *Evaluasi Hasil Belajar*. Edited by Syukrul Hamdi. Lombok: Universitas Hamzanwadi Press, 2017.
- Setiawan, Nugraha. “Teknik Sampling,” 2005, 25–28.
- Sudijono, Anas. *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada, 2012.
- Sudjana, Nana. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya, 2014.
- Sugiyono. *Metode Penelitian Pendidikan Pendekatan Kuantitatif Dan R&D*. Bandung: Alfa Beta, 2010.
- Suharsimi, Arikunto. *Dasar-Dasar Evaluasi Pendidikan*. 2nd ed. Jakarta: Bumi Aksara, 2013.
- Sujarwadi, Sri. “Validitas Dan Reliabilitas Instrumen Penelitian,” 2011, 1–23.
- Wahyuningsih, ErwinTri. “Analisis Butir Soal Tes Objektif Buatan Guru Ulangan

Semester Ganjil Mata Pelajaran Ekonomi Kelas.” Universitas Negeri Yogyakarta, 2015.

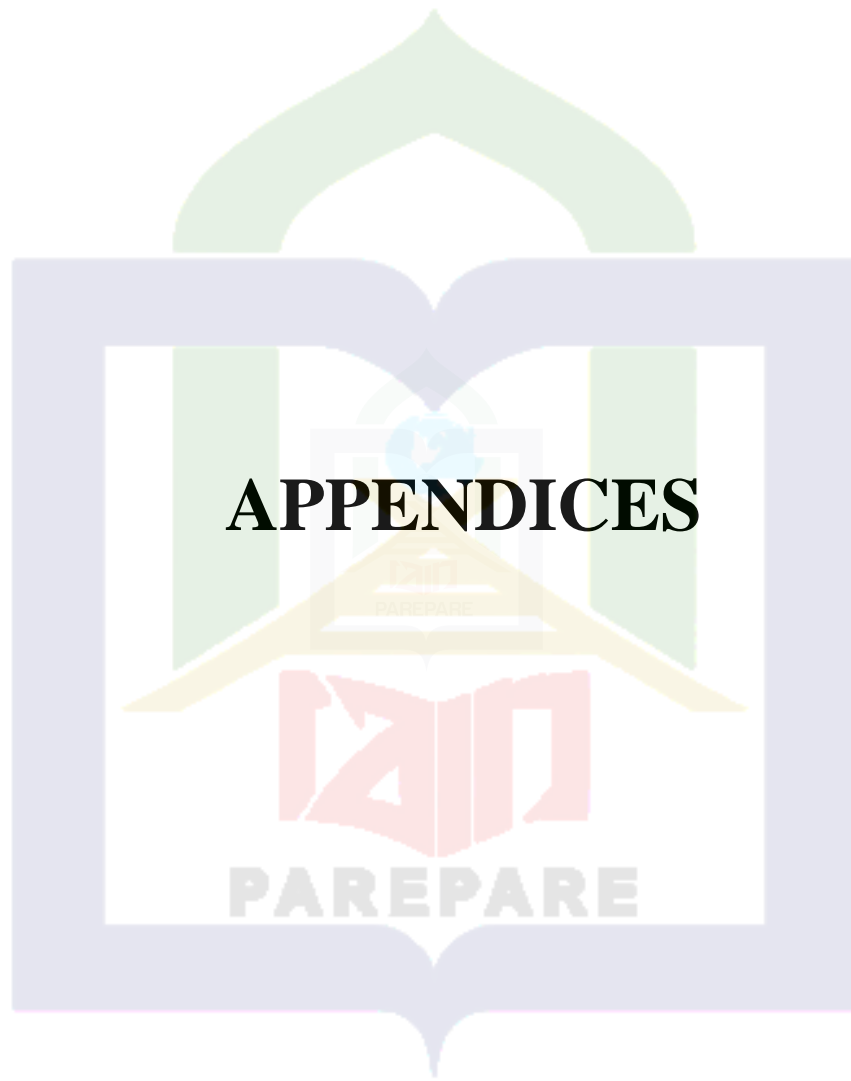
Widiyanto, Joko. *Evaluasi Pembelajaran*. Madiun: Unipma Press, 2018.

Yani, Abet, Ali Fikri Asri, and Ahmad Burhan. “Analisis Tingkat Kesukaran, Daya Pembeda Dan Fungsi Distraktor Soal Ujian Semester Ganjil Mata Pelajaran Produktif Di SMK Negeri 1 Indralaya Utara,” 2013, 98–115.

Zainal, Arifin. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya, 2014.


Zubair, Muhammad Kamal et. all. *Penulisan Karya Ilmiah Berbasis Teknologi Informasi*. Parepare: IAIN Parepare Nusantara Press, 2020





APPENDICES

Appendix 1: Title Certificate


**KEPUTUSAN
DEKAN FAKULTAS TARBIIYAH
NOMOR : 3223 TAHUN 2022
TENTANG
PENETAPAN PEMBIMBING SKRIPSI MAHASISWA FAKULTAS TARBIIYAH
INSTITUT AGAMA ISLAM NEGERI PAREPARE**

DEKAN FAKULTAS TARBIIYAH

Menimbang : a. Bahwa untuk menjamin kualitas skripsi mahasiswa Fakultas Tarbiyah IAIN Parepare, maka dipandang perlu penetapan pembimbing skripsi mahasiswa tahun 2022;
b. Bahwa yang tersebut namanya dalam surat keputusan ini dipandang cakap dan mampu untuk diserahkan tugas sebagai pembimbing skripsi mahasiswa.

Mengingat : 1. Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional;
2. Undang-undang Nomor 12 Tahun 2005 tentang Guru dan Dosen;
3. Undang-undang Nomor 12 Tahun 2012 tentang Pendidikan Tinggi;
4. Peraturan Pemerintah RI Nomor 17 Tahun 2010 tentang Pengelolaan dan Penyelenggaraan Pendidikan;
5. Peraturan Pemerintah RI Nomor 13 Tahun 2015 tentang Perubahan Kedua atas Peraturan Pemerintah RI Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan;
6. Peraturan Presiden RI Nomor 29 Tahun 2018 tentang Institut Agama Islam Negeri Parepare;
7. Keputusan Menteri Agama Nomor 394 Tahun 2003 tentang Pembukaan Program Studi;
8. Keputusan Menteri Agama Nomor 387 Tahun 2004 tentang Petunjuk Pelaksanaan Pembukaan Program Studi pada Perguruan Tinggi Agama Islam;
9. Peraturan Menteri Agama Nomor 35 Tahun 2018 tentang Organisasi dan Tata Kerja IAIN Parepare;
10. Peraturan Menteri Agama Nomor 16 Tahun 2019 tentang Statuta Institut Agama Islam Negeri Parepare.

Memperhatikan : a. Surat Pengesahan Daftar Isian Pelaksanaan Anggaran Petikan Nomor: SP DIPA-025.04.2.307381/2022, tanggal 17 November 2021 tentang DIPA IAIN Parepare Tahun Anggaran 2022;
b. Surat Keputusan Rektor Institut Agama Islam Negeri Parepare Nomor: 494 Tahun 2022, tanggal 31 Maret 2022 tentang Pembimbing Skripsi Mahasiswa Fakultas Tarbiyah IAIN Parepare Tahun 2022.

MEMUTUSKAN

Menetapkan : **KEPUTUSAN DEKAN FAKULTAS TARBIIYAH TENTANG PEMBIMBING SKRIPSI MAHASISWA FAKULTAS TARBIIYAH INSTITUT AGAMA ISLAM NEGERI PAREPARE TAHUN 2022;**


Kesatu : Menunjuk saudara: 1. Dr. Magdhalena Tjalla, M.Hum
2. Yessicka Noviasmy, M.Pd.
Masing-masing sebagai pembimbing utama dan pendamping bagi mahasiswa :
Nama : Risma
NIM : 19.1300.078
Program Studi : Pendidikan Bahasa Inggris
Judul Skripsi : An Analysis of The English Teacher-Made Tests For Second Grade Students At Junior High School 3 Mamuju


Kedua : Tugas pembimbing utama dan pendamping adalah membimbing dan mengarahkan mahasiswa mulai pada penyusunan proposal penelitian sampai menjadi sebuah karya ilmiah yang berkualitas dalam bentuk skripsi;

Ketiga : Segala biaya akibat diterbitkannya surat keputusan ini dibebankan kepada anggaran belanja IAIN Parepare;

Keempat : Surat keputusan ini diberikan kepada masing-masing yang bersangkutan untuk diketahui dan dilaksanakan sebagaimana mestinya.

Ditetapkan di : Parepare
Pada Tanggal : 29 Agustus 2022


Dekan,
I. Zulhah



Appendix 2: Instruments of The Research

	<p>KEMENTERIAN AGAMA REPUBLIK INDONESIA INSTITUT AGAMA ISLAM NEGERI PAREPARE FAKULTAS TARBIYAH Jl.Amal Bakti No. 8 Soreang 91132 Telp. (0421) 21307</p>
<p>VALIDASI INSTRUMENT PENELITIAN PENULISAN SKRIPSI</p>	

Nama Mahasiswa : Risma
 NIM : 19.1300.078
 Fakultas/Prodi : Tarbiyah / Pendidikan Bahasa Inggris
 Judul : An Analysis of the English Teacher-made Test for Second Grade Students at UPT SMKN 3 Parepare

PEDOMAN INSTRUMENT DOKUMENTASI

No	Tentang	Jenis Dokumen
1.	Soal Ujian Akhir Semester Ganjil Kelas XI Multimedia	Data dan Foto
2.	Kunci Jawaban	Data dan Foto
3.	Lembar Jawaban siswa Kelas XI Multimedia 1	Data dan Foto


Parepare, 24th January 2023

Approved by
 Consultant Commissions

Consultant

Co-Consultant


Dr. Magdahatena Tjalla, M.Hum
 NIP. 197003202005012006


Yessicka Noviasmy, M.Pd.
 NIDN. 2029119004

Appendix 3: Research Papers

**KEMENTERIAN AGAMA REPUBLIK INDONESIA**
INSTITUT AGAMA ISLAM NEGERI PAREPARE
FAKULTAS TARBİYAH
Alamat : Jl. Asad Bakri No. 08 Srengeng Parepare 91132 telp 0421) 21307 Fax.24494
PO Box 999 Parepare 91100, website: www.iainparepare.ac.id, email: mail@iainparepare.ac.id

Nomor : B.1979/In.39/FTAR.01/PP.00.9/05/2023 17 Mei 2023
Lampiran : 1 Bundel Proposal Penelitian
Hal : Permohonan Rekomendasi Izin Penelitian

Yth. Kepala Dinas Penanaman Modal dan PTSP
Provinsi Sulawesi Selatan
di,-
Makassar

Assalamu Alaikum Wr. Wb.

Dengan ini disampaikan bahwa mahasiswa Institut Agama Islam Negeri Parepare :

Nama	: Risma
Tempat/Tgl. Lahir	: Mamuju, 8 Desember 2000
NIM	: 19.1300.078
Fakultas / Program Studi	: Tarbiyah/ Pendidikan Bahasa Inggris
Semester	: VIII (Delapan)
Alamat	: Dusun Lingk. Sampaga, Kel. Mamunyu, Kec. Mamuju, Kab. Mamuju, Prov. Sulawesi Barat

Bermaksud akan mengadakan penelitian di wilayah Kota Parepare dalam rangka penyusunan skripsi yang berjudul "An Anlysis Of The English Teacher-Made Test For Second Grade Students At UPT SMKN 3 Parepare". Pelaksanaan penelitian ini direncanakan pada bulan Mei sampai bulan Juni Tahun 2023.
Demikian permohonan ini disampaikan atas perkenaan dan kerjasamanya diucapkan terima kasih.

Wassalamu Alaikum Wr. Wb.


Dekan
Dr. Zulfah, M.Pd.
NIP. 19830420 200801 2 010

Tembusan:

- 1 Rektor IAIN Parepare
- 2 Dekan Fakultas Tarbiyah



PEMERINTAH PROVINSI SULAWESI SELATAN
DINAS PENANAMAN MODAL DAN PELAYANAN TERPADU SATU PINTU

Jl. Bougenville No.5 Telp. (0411) 441077 Fax. (0411) 448936
Website : <http://simap-new.sulselprov.go.id> Email : ptsp@sulselprov.go.id
Makassar 90231

Nomor : 17646/S.01/PTSP/2023 Kepada Yth.
Lampiran : - Dinas Pendidikan Provinsi Sulawesi Selatan
Perihal : izin penelitian

di-
Tempat

Berdasarkan surat Dekan Fak. tarbiyah Institut Agama Islam Negeri Parepare Nomor : B.1979/IN.39/FTAR.01/PP.00.9/05/2023 tanggal 17 Mei 2023 perihal tersebut diatas, mahasiswa/peneliti dibawah ini:

Nama : **RISMA**
Nomor Pokok : 19.1300.078
Program Studi : Pendidikan Bahasa Inggris
Pekerjaan/Lembaga : Mahasiswa (S1)
Alamat : Jl. Amal Bakti No. 8 Soreang, Kota Parepare
PROVINSI SULAWESI SELATAN

Bermaksud untuk melakukan penelitian di daerah/kantor saudara dalam rangka menyusun SKRIPSI, dengan judul :

" AN ANALYSIS OF THE ENGLISH TEACHER-MADE TEST FOR SECOND GRADE STUDENTS AT UPT SMKN 3 PAREPARE "

Yang akan dilaksanakan dari : Tgl. **25 Mei s/d 30 Juni 2023**

Sehubungan dengan hal tersebut diatas, pada prinsipnya kami **menyetujui** kegiatan dimaksud dengan ketentuan yang tertera di belakang surat izin penelitian.

Demikian Surat Keterangan ini diberikan agar dipergunakan sebagaimana mestinya.

Diterbitkan di Makassar
Pada Tanggal 25 Mei 2023

A.n. GUBERNUR SULAWESI SELATAN
PLT. KEPALA DINAS PENANAMAN MODAL DAN PELAYANAN TERPADU
SATU PINTU PROVINSI SULAWESI SELATAN



Drs. MUH SALEH, M.Si.
Pangkat : PEMBINA UTAMA MUDA
Nip : 19690717 199112 1002

Tembusan Yth
1. Dekan Fak. tarbiyah Institut Agama Islam Negeri Parepare;
2. *Pertinggal.*

Nomor: 17646/S.01/PTSP/2023

KETENTUAN PEMEGANG IZIN PENELITIAN :

1. Sebelum dan sesudah melaksanakan kegiatan, kepada yang bersangkutan melapor kepada Bupati/Walikota C q. Kepala Bappelitbangda Prov. Sulsel, apabila kegiatan dilaksanakan di Kab/Kota
2. Penelitian tidak menyimpang dari izin yang diberikan
3. Mentaati semua peraturan perundang-undangan yang berlaku dan mengindahkan adat istiadat setempat
4. Menyerahkan 1 (satu) eksamplar hardcopy dan softcopy kepada Gubernur Sulsel. Cq. Kepala Badan Perencanaan Pembangunan Penelitian dan Pengembangan Daerah Prov. Sulsel
5. Surat izin akan dicabut kembali dan dinyatakan tidak berlaku apabila ternyata pemegang surat izin ini tidak mentaati ketentuan tersebut diatas.

REGISTRASI ONLINE IZIN PENELITIAN DI WEBSITE :
<https://izin-penelitian.sulselprov.go.id>





PEMERINTAH PROPINSI SULAWESI SELATAN
DINAS PENDIDIKAN
UPT SMKN 3 PAREPARE

Jl. Karaeng Burane No. 16 Tlp/Fax (0421) 21266 Kota Parepare 91111
Email : smkntiga.parepare@gmail.com/ Website : www.smkntigaparepare.sch.id

SURAT KETERANGAN

Nomor : 421.5/148 - UPT SMKN.3/PARE/DISDIK

Yang bertanda tangan dibawah ini, Kepala UPT SMKN 3 Parepare menerangkan bahwa :

Nama : RISMA
NIM : 19.1300.078
Program Studi : PENDIDIKAN BAHASA INGGRIS
Perguruan Tinggi : IAIN PAREPARE

Benar telah melakukan penelitian di UPT SMKN 3 Parepare untuk memperoleh data yang berkaitan dengan penyusunan Skripsi yang berjudul : **"AN ANALYSIS OF THE ENGLISH TEACHER-MADE TEST FOR SECOND GRADE STUDENTS AT UPT SMKN 3 PAREPARE"** .

Demikian Surat Keterangan ini diberikan kepada yang bersangkutan untuk dilaksanakan sebagaimana mestinya.

Parepare, 27 Juni 2023

Kepala UPT SMKN 3 Parepare,



ANDI RAEHANA R. S.Pd.MM

Pangsa Pembina Tk. I

NIP : 19721102 200005 2 001



#BerAKHLAK
#SIPAKATAU

#CERDASKI!
Cekatan BerEtika Berintegritas Berdedikasi
Amanah Bertanggungjawab Inovatif

SETULUS RATI, SEPENUH JIWA, SEKUAT SAGA
PENGENDAH SULAWESI SELATAN

Appendix 4: Final Test For Second Grade at UPT SMKN 3 Parepare

**SOAL ULANGAN SEMESTER GANJIL
TAHUN PELAJARAN 2022/2023**

Mata Pelajaran : BAHASA INGGRIS
Kelas : XI

Choose the best answer by crossing A, B, C, D, or E on your answer sheet!

1. Jade : What do you think of my drawing?
Rose : It's amazing, but I think you should erase the scribbles over here.
Jade : Thank you so much for you opinion, Rose!
Rose : My pleasure.
From the dialogue above, Rose is...

A. Giving help
B. Asking a question
C. Giving an opinion
D. Asking for attention
E. Asking for help

2. Bowo : I feel tired and I feel dizzy.
Sri : I think Don't leave the bed if it is not necessary.

A. You must sing
B. I will take you to the hospital
C. I must take some rest, too
D. You should lie down and have some rest
E. You can see the doctor tonight.

3. Eric : I think our city is very hot at the moment.
Era : I don't think so Our city is much cooler than other cities in this country.

A. I know it
B. I am thinking of

- C. He forget it
- D. See you
- E. In my opinion

4. Layla : Thanks for meeting with me during your lunch hour. I appreciate it.
 Monica : No problem. I'm happy to help. What's going on?
 Layla : Oh you know, the usual. Should I take this new job? Or do I stick with my current one?

Monica : Well, I think it's time for a change, don't you? They pay you late and you are unhappy.

Layla : Do you really think so?

Monica : I know so. And I've been listening to you complain for over a year now. Trust me. Take the job. What do you have to lose?

What does Monica suggest for Layla?

- A. accept the offer for her new job
 - B. have same job with her
 - C. join her job in her office
 - D. change Layla's opinion
 - E. reject the offer for her new job
5. Dave : So what do you think of my singing?
 John : It's really good, but I suggest trying singing in a high tune.
 Dave : Thanks, John.
 John : No problem, Dave!
- From the dialogue above, John is...
- A. Asking for an opinion
 - B. Giving an opinion
 - C. Giving help
 - D. Giving attention
 - E. Asking for help

6. Man : Hi, Stephany. I hear that you're getting an English test tomorrow. I think you should prepare it well.

Woman : Yes, that's right, Bob. I have been learning all day long.

Man : That's very good. I hope you will get a good score.

Woman : I hope so. Thanks Bob.

Narrator : What are they talking about?

- A. Learning English
- B. Expressing hope
- C. English Test preparation
- D. Good score
- E. Getting an English score

7. Waiter : Good morning ,?

Guest : I'd like to start with a glass of hot coffee ,please

What is the right expression to complete the dialog above?

- A. Are you ready to order, sir ?
- B. What would you like to drink, sir?
- C. How are you, sir ?
- D. Have you made a reservation before, sir?
- E. Are you OK sir?

8. Amran :What's your opinion of the plans for the new Officer ?

Riska :

- A. I feel that your opinion is right
- B. We don't have any opinion now.
- C. We'll have a planning meeting.
- D. We should meeting
- E. I don't really care.

9. Receptionist : Good morning, reception. May I help you?

Mr.Lungga : Yes, I have to check out at 06.30 tomorrow morning May I have

breakfast at six o'clock?

Receptionist : Don't worry,....

- A. We can't prepare so early
- B. Our hotel has many waiters
- C. Could you mention the exact time ?
- D. We will prepare it for you on time
- E. I can't help you, sir

10. Reski : Excuse me, can I help you, sir?

Andi : Yes, Can you show me where is the manager office?

Reski :

- A. If I don't care about you
- B. If you follow me, I would tell you the office.
- C. If you passed the hall, you would find it.
- D. If you go there on foot, you will find it.
- E. If you follow me, I'll show you the way



Appendix 5 : The Result Calculation For Test Students

NO.	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
1	Elif Hartiningsih	1	1	1	1	1	1	1	0	1	1	9
2	Aisyah Fadhillah Muh. Azis	1	1	1	1	1	1	1	0	1	1	9
3	Astri Cindy Manchetti	1	1	1	1	1	1	1	0	1	1	9
4	Andi Mahathir Muhammad	1	1	1	1	1	1	1	0	1	0	8
5	Ayu Dwi Lestari	1	1	1	0	1	1	1	0	1	1	8
6	Amelya Nanda	1	1	1	1	1	0	1	0	1	1	8
7	Aisyah Atirah	1	1	1	1	1	1	0	0	1	1	8
8	Chelsea Olivia Harum	1	1	0	1	1	1	1	0	1	1	8
9	Dirham Dwi Yulianto Fereira	1	1	0	0	1	1	1	0	1	1	7
10	Fito Anugrah Nurzaman	1	1	1	1	1	0	0	0	1	1	7
11	Airin Tenriani	1	1	0	0	1	1	1	0	1	1	7
12	Bania Nur Sakinah	1	1	0	1	0	1	1	0	1	1	7
13	Akhiryati Berlian	1	0	1	1	1	1	0	1	1	0	7
14	Audya Pertiwi	1	0	1	1	1	0	1	1	0	1	7
15	Dian Sardi Jay	1	1	0	0	1	1	1	1	1	0	7
16	Azyurah Dwi Putri	1	1	1	0	0	1	0	1	1	1	7
17	Alfiana Muhsin	1	0	1	1	1	0	1	0	0	1	6
18	Alya Sephira Fitriani	1	1	1	1	1	0	0	0	1	0	6
19	Citra Angreni	1	1	1	0	1	1	0	0	1	0	6
20	Acyuta Mujahid. S	1	1	1	1	1	1	0	0	0	0	6
21	Farah Quin Afriyani	1	0	1	0	1	0	1	1	0	1	6
22	Astri	1	0	0	1	1	1	0	0	1	0	5
23	Aulia Rahman	0	1	0	1	1	1	1	0	0	0	5
24	Deby Grace Mandri	0	0	1	0	1	1	1	1	0	0	5
25	Indah Suci Amelia	1	0	0	1	1	0	0	0	1	0	4
26	Aisyah Ramadhani	0	1	0	0	1	1	0	1	0	0	4
27	Emi	1	0	0	1	1	0	1	0	0	0	4
28	Ayu Andini	0	1	0	1	0	1	0	0	0	0	3
29	Arnita Melani	1	0	0	0	0	0	0	1	0	0	2
30	Adam Subarkah	0	1	0	0	1	0	0	0	0	0	2
31	Febriyani Nurul Aini Pipin	0	0	0	0	0	0	1	0	0	1	2

NO.	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
32	Andini Pertiwi	0	0	0	1	0	0	0	0	0	0	1
33	Delia Istiqinah	0	0	0	0	0	0	0	0	0	1	1
	TOTAL	25	21	17	20	26	20	18	8	19	17	191

UPPER GROUP

NO	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
1	Elif Hartiningsih	1	1	1	1	1	1	1	0	1	1	9
2	Aisyah Fadhillah Muhammad Azis	1	1	1	1	1	1	1	0	1	1	9
3	Astri Cindy Manchetti	1	1	1	1	1	1	1	0	1	1	9
4	Andi Mahathir Muhammad	1	1	1	1	1	1	1	0	1	0	8
5	Ayu Dwi Lestari	1	1	1	0	1	1	1	0	1	1	8
6	Amelya Nanda	1	1	1	1	1	0	1	0	1	1	8
7	Aisyah Atirah	1	1	1	1	1	1	0	0	1	1	8
8	Chelsea Olivia Harum	1	1	0	1	1	1	1	0	1	1	8
9	Dirham Dwi Yulianto Ferreira	1	1	0	0	1	1	1	0	1	1	7
	TOTAL	9	9	7	7	9	8	8	0	9	8	

LOWER GROUP

NO	NAME	ITEM										TOTAL
		1	2	3	4	5	6	7	8	9	10	
1	Indah Suci Amelia	1	0	0	1	1	0	0	0	1	0	4
2	Aisyah Ramadhani	0	1	0	0	1	1	0	1	0	0	4
3	Emi	1	0	0	1	1	0	1	0	0	0	4
4	Ayu Andini	0	1	0	1	0	1	0	0	0	0	3
5	Arnita Melani	1	0	0	0	0	0	0	1	0	0	2
6	Adam Subarkah	0	1	0	0	1	0	0	0	0	0	2
7	Febriayani Nurul Aini Pipin	0	0	0	0	0	0	1	0	0	1	2
8	Andini Pertiwi	0	0	0	1	0	0	0	0	0	0	1
9	Delia Istiqinah	0	0	0	0	0	0	0	0	0	1	1
	TOTAL	3	3	0	4	4	2	2	2	1	2	

A. VALIDITY

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Where:

r_{pbi} = Biserial point correlation

M_p = Mean of students who answered correctly

M_t = Total mean score

S_t = Total standard deviation

p = Proportion of correct answers to all answers

q = Proportion of incorrect answers to all answers

$r_{pbis} < 0.3$ = invalid

$r_{pbis} \geq 0.3$ = valid

1. Menghitung M_p

$$\begin{aligned} \text{➤ MP NO. 1} &= \frac{9+9+9+8+8+8+8+8+7+7+7+7+7+7+7+6+6+6+6+6+5+4+4+2}{25} \\ &= \frac{168}{25} = 6,72 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 2} &= \frac{9+9+9+8+8+8+8+8+7+7+7+7+7+7+6+6+6+5+4+3+2}{21} \\ &= \frac{141}{21} = 6,71 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 3} &= \frac{9+9+9+8+8+8+8+7+7+7+7+6+6+6+6+6+5}{17} \\ &= \frac{122}{17} = 6,18 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 4} &= \frac{9+9+9+8+8+8+8+7+7+7+7+6+6+6+5+5+4+4+3+1}{20} \\ &= \frac{127}{20} = 6,35 \end{aligned}$$

$$\begin{aligned} \text{➤ MP NO. 5} &= \frac{9+9+9+8+8+8+8+8+7+7+7+7+7+7+6+6+6+6+6+5+5+5+4+4+4+2}{26} \\ &= \frac{168}{26} = 6,46 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 6} &= \frac{9+9+9+8+8+8+8+7+7+7+7+7+7+6+6+5+5+5+4+3}{20} \\ &= \frac{135}{20} = 6,75 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 7} &= \frac{9+9+9+8+8+8+8+7+7+7+7+7+6+6+5+5+4+2}{18} \\ &= \frac{122}{18} = 6,78 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 8} &= \frac{7+7+7+7+6+6+5+4}{8} \\ &= \frac{49}{8} = 6,125 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 9} &= \frac{9+9+9+8+8+8+8+8+7+7+7+7+7+7+7+6+6+5+4}{19} \\ &= \frac{137}{19} = 7,21 \end{aligned}$$

$$\begin{aligned} \text{➤ Mp NO. 10} &= \frac{9+9+9+8+8+8+8+7+7+7+7+7+7+6+6+2+1}{17} \\ &= \frac{116}{17} = 6,82 \end{aligned}$$

2. Menghitung Mt (Mean Total)

$$= \frac{9+9+9+8+8+8+8+8+7+7+7+7+7+7+7+7+6+6+6+6+6+5+5+5+4+4+4+3+2+2+2+1+1}{33}$$

$$\text{Mt} = \frac{191}{33} = 5,79$$

3. Menghitung St

No. _____
Date : _____

3. Menghitung S_t $\rightarrow 5,79$

$$S_t = \sqrt{\frac{\sum (m - m_t)^2}{n-1}}$$

$$S_t = \sqrt{\frac{(9-5,79)^2 + (9-5,79)^2 + (9-5,79)^2 + (8-5,79)^2 + (8-5,79)^2 + (8-5,79)^2 + (8-5,79)^2 + (8-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (7-5,79)^2 + (6-5,79)^2 + (6-5,79)^2 + (6-5,79)^2 + (6-5,79)^2 + (6-5,79)^2 + (6-5,79)^2 + (5-5,79)^2 + (5-5,79)^2 + (5-5,79)^2 + (4-5,79)^2 + (4-5,79)^2 + (4-5,79)^2 + (3-5,79)^2 + (2-5,79)^2 + (2-5,79)^2 + (2-5,79)^2 + (1-5,79)^2 + (1-5,79)^2}{33-1}}$$

$$= \sqrt{\frac{(3,21)^2 + (3,21)^2 + (3,21)^2 + (2,21)^2 + (2,21)^2 + (2,21)^2 + (2,21)^2 + (2,21)^2 + (1,21)^2 + (1,21)^2 + (1,21)^2 + (1,21)^2 + (1,21)^2 + (1,21)^2 + (0,21)^2 + (0,21)^2 + (0,21)^2 + (0,21)^2 + (0,21)^2 + (0,21)^2 + (-0,79)^2 + (-0,79)^2 + (-0,79)^2 + (-0,79)^2 + (-1,79)^2 + (-1,79)^2 + (-1,79)^2 + (-2,79)^2 + (-2,79)^2 + (-2,79)^2 + (-2,79)^2 + (-3,79)^2 + (-3,79)^2 + (-3,79)^2 + (-3,79)^2 + (-4,79)^2 + (-4,79)^2}{32}}$$

$$= \sqrt{\frac{10,30 + 10,30 + 10,30 + 4,88 + 4,88 + 4,88 + 4,88 + 4,88 + 1,46 + 1,46 + 1,46 + 1,46 + 1,46 + 1,46 + 0,09 + 0,09 + 0,09 + 0,09 + 0,09 + 0,09 + 0,62 + 0,62 + 0,62 + 3,20 + 3,20 + 3,20 + 7,78 + 14,36 + 14,36 + 14,36 + 22,94 + 22,94}{32}}$$

$$= \sqrt{\frac{175,5}{32}}$$

$$= \sqrt{5,48}$$

$$S_t = 2,341$$

PAREPARE

4. Menghitung p dan q

$$1. \quad p = \frac{25}{33} = 0,8 \quad \rightarrow \quad pq = 0,8 \times 0,6 \\ = 0,16$$

$$q = 1 - p \\ = 1 - 0,8 = 0,2$$

$$2. \quad p = \frac{21}{33} = 0,6 \quad \rightarrow \quad pq = 0,6 \times 0,4 \\ = 0,24$$

$$q = 1 - p \\ = 1 - 0,6 = 0,4$$

$$3. \quad p = \frac{17}{33} = 0,5 \quad \rightarrow \quad pq = 0,5 \times 0,5 \\ = 0,25$$

$$q = 1 - p \\ = 1 - 0,5 = 0,5$$

$$4. \quad p = \frac{20}{33} = 0,6 \quad \rightarrow \quad pq = 0,6 \times 0,4 \\ = 0,24$$

$$q = 1 - p \\ = 1 - 0,6 = 0,4$$

$$5. \quad p = \frac{26}{33} = 0,8 \quad \rightarrow \quad pq = 0,8 \times 0,2 \\ = 0,16$$

$$q = 1 - p \\ = 1 - 0,8 = 0,2$$

$$6. \quad p = \frac{20}{33} = 0,6$$

$$q = 1 - p$$

$$= 1 - 0,6 = 0,4$$

$$pq = 0,6 \times 0,4$$

$$= 0,24$$

$$7. \quad p = \frac{18}{33} = 0,5$$

$$q = 1 - p$$

$$= 1 - 0,5 = 0,5$$

$$pq = 0,5 \times 0,5$$

$$= 0,25$$

$$8. \quad p = \frac{8}{33} = 0,2$$

$$q = 1 - p$$

$$= 1 - 0,2 = 0,8$$

$$pq = 0,2 \times 0,8$$

$$= 0,16$$

$$9. \quad p = \frac{19}{33} = 0,6$$

$$q = 1 - p$$

$$= 1 - 0,6 = 0,4$$

$$pq = 0,6 \times 0,4$$

$$= 0,24$$

$$10. \quad p = \frac{17}{33} = 0,5$$

$$q = 1 - p$$

$$= 1 - 0,5 = 0,5$$

$$pq = 0,5 \times 0,5$$

$$= 0,25$$

5. Menghitung Validity

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 1} &= \frac{6,72-5,79}{2,34} \sqrt{\frac{0,8}{0,2}} \\ &= \frac{0,93}{2,34} \times 2 = \frac{1,86}{2,34} = 0,79 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 2} &= \frac{6,71-5,79}{2,34} \sqrt{\frac{0,6}{0,4}} \\ &= \frac{0,92}{2,34} \times 1,2 = \frac{1,10}{2,34} = 0,47 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 3} &= \frac{7,18-5,79}{2,34} \sqrt{\frac{0,5}{0,5}} \\ &= \frac{1,39}{2,34} \times 1 = \frac{1,39}{2,34} = 0,59 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 4} &= \frac{6,35-5,79}{2,34} \sqrt{\frac{0,6}{0,4}} \\ &= \frac{0,56}{2,34} \times 1,2 = \frac{0,67}{2,34} = 0,29 \text{ (TIDAK VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 5} &= \frac{6,46-5,79}{2,34} \sqrt{\frac{0,8}{0,2}} \\ &= \frac{0,67}{2,34} \times 2 = \frac{1,34}{2,34} = 0,57 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 6} &= \frac{6,75-5,79}{2,34} \sqrt{\frac{0,6}{0,4}} \\ &= \frac{0,96}{2,34} \times 1,2 = \frac{1,15}{2,34} = 0,49 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 7} &= \frac{6,78-5,79}{2,34} \sqrt{\frac{0,5}{0,5}} \\ &= \frac{0,99}{2,34} \times 1 = \frac{0,99}{2,34} = 0,42 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 8} &= \frac{6,12-5,79}{2,34} \sqrt{\frac{0,2}{0,8}} \\ &= \frac{0,33}{2,34} \times 0,5 = \frac{0,16}{2,34} = 0,07 \text{ (TIDAK VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 9} &= \frac{7,21-5,79}{2,34} \sqrt{\frac{0,6}{0,4}} \\ &= \frac{1,42}{2,34} \times 1,2 = \frac{1,70}{2,34} = 0,73 \text{ (VALID)} \end{aligned}$$

$$\begin{aligned} \text{➤ } r_{pbi} \text{ NO. 10} &= \frac{6,82-5,79}{2,34} \sqrt{\frac{0,5}{0,5}} \\ &= \frac{1,03}{2,34} \times 1 = \frac{1,03}{2,34} = 0,44 \text{ (VALID)} \end{aligned}$$

INVALID = NO. 4 and 8

VALID = NO. 1, 2, 3, 5, 6, 7, 9, and 10

B. RELIABILITY

$$KR - 20 = \frac{n}{n-1} \left(\frac{S_t^2 - \sum pq}{S_t^2} \right)$$

Diketahui:

- $n = 33$
- $st^2 = 2,34^2$
 $= 5,48$
- $\sum pq = 0,16 + 0,24 + 0,25 + 0,24 + 0,16 + 0,24 + 0,25 + 0,16 + 0,24 + 0,25$
 $= 2,19$

$$KR - 20 = \frac{33}{33-1} \left(\frac{5,48 - 2,19}{5,48} \right)$$

$$KR - 20 = \frac{33}{32} \left(\frac{3,29}{5,48} \right)$$

$$KR - 20 = 1,03 (0,60)$$

$$KR - 20 = 0,6 \text{ (UNRELIABLE)}$$

C. LEVEL OF DIFFICULTY

$$P = \frac{B}{JS}$$

Where:

P = level of difficulty

B = the number of students who answer the items correctly

JS = number of students in each group

- P NO. 1 = $\frac{9+3}{9+9} = \frac{12}{18} = 0,67$ (MODERATE)
- P NO. 2 = $\frac{9+3}{9+9} = \frac{12}{18} = 0,67$ (MODERATE)
- P NO. 3 = $\frac{7+0}{9+9} = \frac{7}{18} = 0,38$ (MODERATE)
- P NO. 4 = $\frac{7+4}{9+9} = \frac{11}{18} = 0,61$ (MODERATE)
- P NO. 5 = $\frac{9+4}{9+9} = \frac{13}{18} = 0,72$ (EASY)
- P NO. 6 = $\frac{8+2}{9+9} = \frac{10}{18} = 0,55$ (MODERATE)
- P NO. 7 = $\frac{8+2}{9+9} = \frac{10}{18} = 0,55$ (MODERATE)
- P NO. 8 = $\frac{0+2}{9+9} = \frac{2}{18} = 0,11$ (DIFFICULTY)

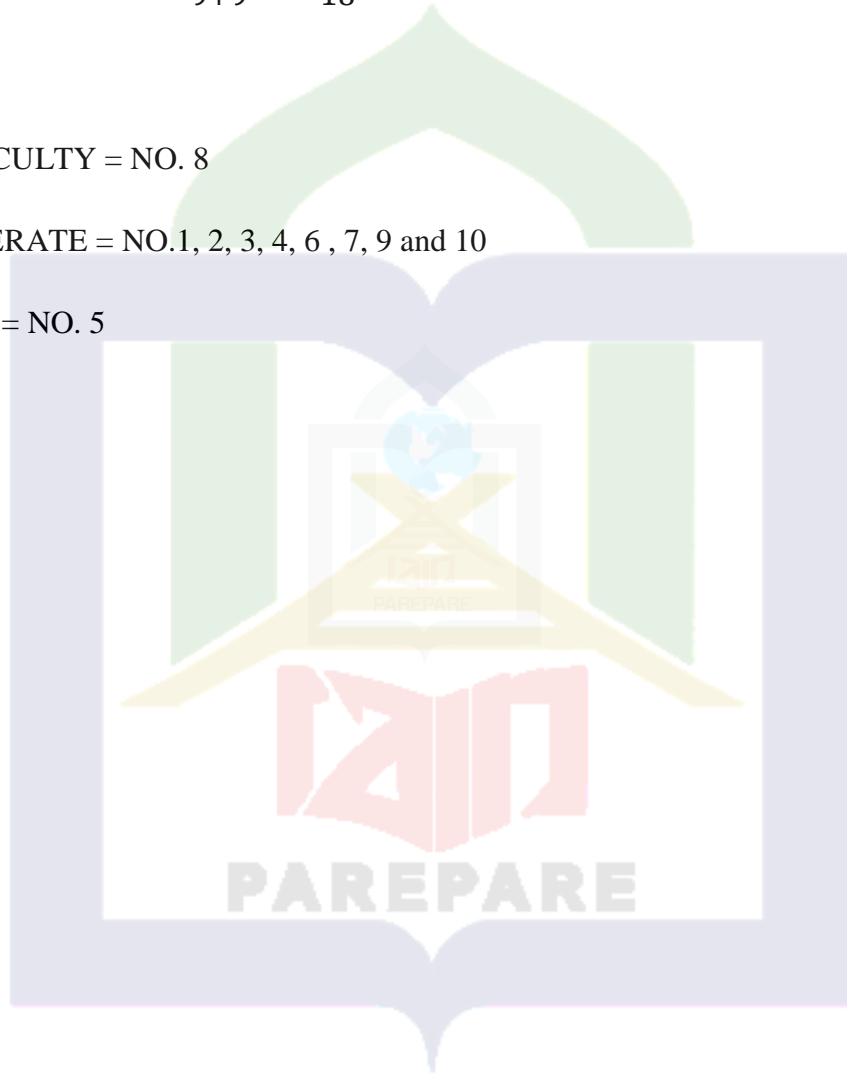
➤ P NO. 9 = $\frac{9+1}{9+9} = \frac{10}{18} = 0,55$ (MODERATE)

➤ P NO. 10 = $\frac{8+2}{9+9} = \frac{10}{18} = 0,55$ (MODERATE)

DIFFICULTY = NO. 8

MODERATE = NO.1, 2, 3, 4, 6, 7, 9 and 10

EASY = NO. 5



D. DICRIMINATION INDEX

$$DP = \frac{WL - WH}{n}$$

Where:

DP = Discriminating Power

WL = the number of students in the lower group who answer the items incorrectly

WH = the number of students in the upper group who answer the items incorrectly

n = 27 % X N

$$\text{➤ DP NO. 1} = \frac{6-0}{9} = \frac{6}{9} = 0,67 \text{ (GOOD)}$$

$$\text{➤ DP NO. 2} = \frac{6-0}{9} = \frac{6}{9} = 0,67 \text{ (GOOD)}$$

$$\text{➤ DP NO. 3} = \frac{9-2}{9} = \frac{7}{9} = 0,78 \text{ (EXCELLENT)}$$

$$\text{➤ DP NO. 4} = \frac{5-2}{9} = \frac{3}{9} = 0,33 \text{ (SATISFACTORY)}$$

$$\text{➤ DP NO. 5} = \frac{5-0}{9} = \frac{5}{9} = 0,55 \text{ (GOOD)}$$

$$\text{➤ DP NO. 6} = \frac{7-1}{9} = \frac{6}{9} = 0,67 \text{ (GOOD)}$$

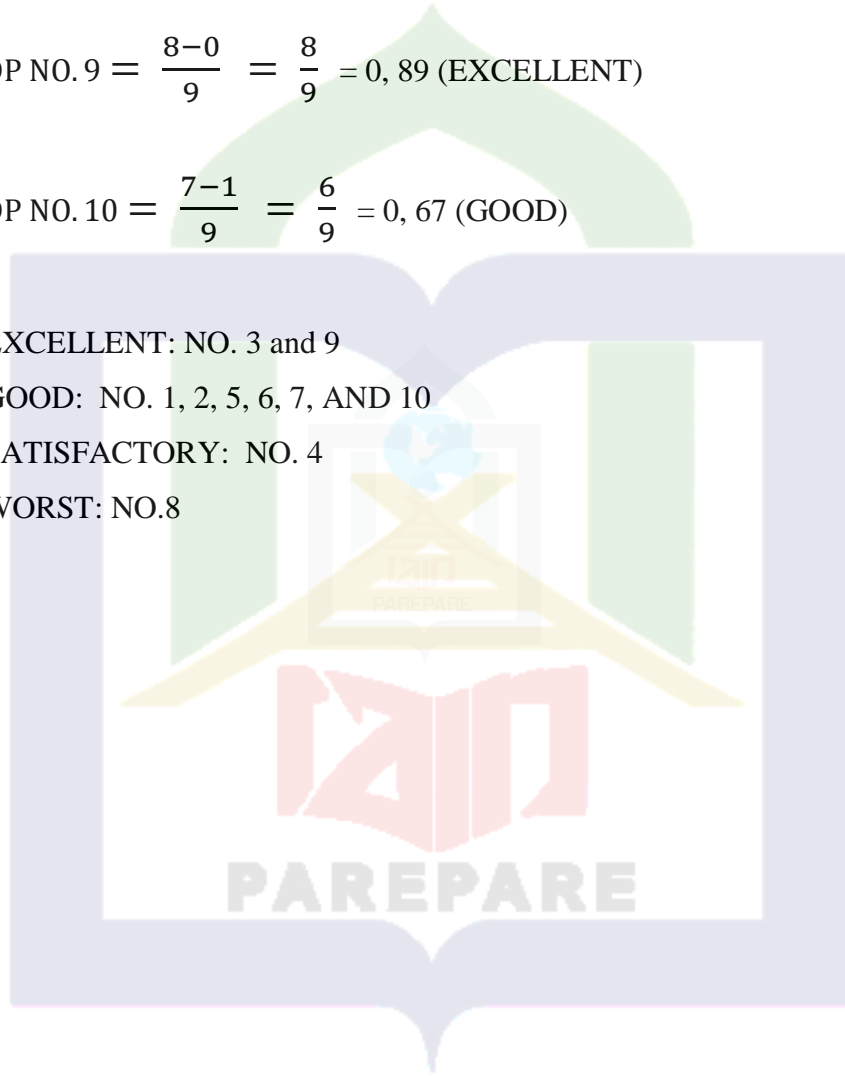
- DP NO. 7 = $\frac{7-1}{9} = \frac{6}{9} = 0,67$ (GOOD)
- DP NO. 8 = $\frac{7-9}{9} = \frac{-2}{9} = -0,22$ (WORST)
- DP NO. 9 = $\frac{8-0}{9} = \frac{8}{9} = 0,89$ (EXCELLENT)
- DP NO. 10 = $\frac{7-1}{9} = \frac{6}{9} = 0,67$ (GOOD)

EXCELLENT: NO. 3 and 9

GOOD: NO. 1, 2, 5, 6, 7, AND 10

SATISFACTORY: NO. 4

WORST: NO.8



E. DISTRACTORS

Item	Option					Total
	A	B	C	D	E	
1	3	2	*25	3	0	33
2	1	6	3	*21	2	33
3	3	10	1	2	*17	33
4	*20	1	1	11	0	33
5	3	*26	1	2	1	33
6	7	2	*20	2	2	33
7	13	*18	0	2	0	33
8	8	*8	11	4	2	33
9	0	5	3	*19	6	33
10	4	4	6	2	*17	33

CURRICULUM VITAE



RISMA, the writer was born December 08th 2000 in Mamuju. Her father's name is Usman T and her mother's name is Isa. She is the youngest of five children, she has two sister and two brother. She began her study in Elementary school at SD Inpres Binanga II kalubibing and graduated on 2013. In the same year, she continued her study to SMPN 3 Mamuju and graduated on 2016. She decided to continue her study to SMAN 1 Mamuju and graduated on 2019. However, she continued her study at State Islamic Institute of Parepare on 2019. During she study in IAIN Parepare and took English Departement as her major. She completed her thesis in the title "An Analysis of English Teacher-Made Test For Second Grade Students at UPT SMKN 3 Parepare"

