

**TEACHERS' PROBLEMS IN DESIGNING SUMMATIVE
TEST AT SENIOR HIGH SCHOOL KEPAHANG**

THESIS

**This thesis is submitted to fulfill the
requirement for 'Sarjana' degree English
Study Program.**



By :

ANGGUN PUTRI SUNDARI

NIM.15551005

**ENGLISH TADRIS STUDY PROGRAM
FACULTY OF TARBIYAH
INSTITUTE COLLEGE OF ISLAMIC STUDIES
(IAIN) CURUP
2022**

SUPERVISORS' APPROVAL

Hal : Pengajuan Skripsi

Kepada

Yth. Bapak Dekan Tarbiyah IAIN Curup

Di

Curup

Assalamu'alaikum Wr.Wb

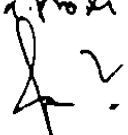
Setelah mengadakan pemeriksaan dan perbaikan seperlunya maka kami berpendapat bahwa skripsi atas nama **Anggun Putri Sundari (15551005)** mahasiswa IAIN Curup Prodi Tadris Bahasa Inggris, yang berjudul **"Teachers' Problems in Designing Summative Test at Senior High School Kepahiang"** sudah dapat diajukan dalam sidang munaqasah Institut Agama Islam Negeri (IAIN) Curup.

Demikian permohonan ini kami ajukan, agar dapat diterima terlebih dahulu diucapkan terima kasih.

Wassalamu'alaikum, Wr.Wb

Curup, June 2022

Advisor

An. Lg. Prodi TB 1


Desfitranita, M.Pd
NIP. 19791225 200912 2 002

Co-Advisor



Henny Septia Utami, M.Pd
NIDN. 2016098903



**KEMENTERIAN AGAMA REPUBLIK INDONESIA
INSTITUT AGAMA ISLAM NEGERI (IAIN) CURUP
FAKULTAS TARBIYAH**

Jl. Dr. Ak Gani No. 01 Kotak Pos 108 Telp
(0732) 2101102179 Fax

Homepage: <http://www.iaincurup.ac.id> Email: admint@iaincurup.ac.id Pos 39119

A P P R O V A L

Nomor: 1515/In.34/F.T/PP.00.9/08/2022

Name : **Anggun Putri Sundari**
NIM : **15551005**
Faculty : **Tarbiyah**
Department : **Tadris Bahasa Inggris**
Title : **Theachers' Problems in Designing English Summative Test at Senior High School Kepahiang**

Has been examined by examining board of the English Study Program of Institut Agama Islam Negeri (IAIN) Curup on :

Day / Date : **Friday, July 29th 2022**
Time : **15:00 Pm-16:30 pm**
At : **Munqasyah Room IV IAIN Curup**

Has been received to fulfill partial requirements for the degree of Strata 1 in English Study Program of Tarbiyah Department of IAIN Curup.

EXAMINERS :

Head

An. *[Signature]* Pro Di TB1

Desfitranita, M. Pd
NIP. 197912252009122002

Secretary

[Signature]

Henny Septia Utami, M. Pd
NIDN. 2016098903

Examiner I

[Signature]
Dr. Sakut Anshori, M. Hum
NIP. 198110202006041002

Examiner II,

[Signature]
Sarwo Edli M. Pd
NIDN. 2020108101

Dean of Faculty Tarbiyah



Dr. H. Hamengkubuwono, M.Pd
NIP. 19650826 199903 1 001

THE STATEMENT OF OWNERSHIP

The writer who sign below:

Name : Anggun Putri Sundari

NIM : 15551005

Fakultas : Tarbiyah

Prodi : TBI (Tadris Bahasa Inggris)

State that the thesis is entitled **“Teachers' Problems in Designing Summative Test at Senior High School Kepahiang”**. This statement is made truly, if in the next day there any mistake the writer ready to accept the punishment or other criticism from IAIN suitable with its regulation.

Curup, June 2022

Writer



ANGGUN PUTRI SUNDARI
NIM. 15551005

PREFACE

All praises to Allah SWT that the writer had finally finished writing her thesis entitled “**Teachers' Problems in Designing Summative Test at Senior High School Kepahiang**”.

This thesis is submitted as a part of the completion for undergraduate degree of strata 1 (S1) in English Study Program of IAIN Curup. The writer realize that this thesis is far being perfect, therefore the writer really appreciates some suggestion and critics for being perfect in the future.

Last but not least, the writer hopes that this thesis will be really useful to those who are interested in this field of study.

Curup, January 2022

Writer



ANGGUN PUTRI SUNDARI
NIM.15551005

ACKNOWLEDGEMENT



Assalamu'alaikum Warrahmatulahi Wabarrakatu.

Alhamdulillah rabbil'amin, all of praise to Allah SWT the all mighty and merciful god who has given the researcher mercy, blessing and guidance so the researcher can finish this research completely, peace and salutation always be given to Prophet Muhammad (Peace Be Upon Him), and all of his family and followers who have been a good example for every muslim in this world and has brought us from the darkness (Jahiliyah) to the lightness (Islam) as we felt together.

The researcher finished this thesis entitled **“Teachers' Problems in Designing Summative Test at Senior High School Kepahiang”**. This thesis is presented in order to fulfill of the requirement for the degree of strata in English study program IAIN Curup. In the process of conducting this research, the researcher accepted valuable support, assistance, guidance, contribution and motivation from others, because of those reason, the writer would like to presents her deepest appreciation to :

1. Mr. Prof. Dr. Idi Warsah, M.Pd.I as the rector of IAIN Curup.
2. Mr. Dr. H. Hamengkubuwono, M.Pd as Dekan of Education (Tarbiyah) Faculty.
3. Mrs. Jumatul Hidayah as the head of English Study Program of IAIN Curup thank you for your support mam.

4. Special thanks for my parents, my father a hero of my life Mr Sundar Mawan and my mom a angel of my heart Mrs Haya Lela Wati, my brother Yudahari Sandi, Prayoga Saputra, Yusrialtaf Musyaffa and my lovely sister Ema Romantika.
5. Mrs Desfitranita, M.Pd as my advisor and also my thesis advisor thank you for your support, guidance since I decided my thesis's title. Allah blesses every step in your life InsyaAllah.
6. Miss Henny Septia Utami, M.Pd as my Co-Advisor, who always give hard words to me then I could learn how to be strong and how to finish this thesis. For all time Allah will give his bless for you. Aamiin
7. My lectures, (Mr Rully Morgana, Mr Sakut Ansori, Mr Sarwo Edy, Mr Paidi Gusmuliana, Mrs Eka Apriani, Mr Bayu Senjahari, Mrs Desfirnita, Mr Benni Azwar, etc) who have taught me many things in class.

Finally, the writer need construction suggestion for being perfect in the future. Hopefully, the result of this research can give beneficial contributionto the development of education in other school.for acknowledgment above, and those are not mentioned, may Allah give them the best reward. Aamiin.

Curup, June 2022

Writer



ANGGUN PUTRI SUNDARI
NIM.15551005

ABSTRACT

Sundari, Anggun Putri. 2022. Teachers' Problems in Designing Summative Test at (Senior High School Kepahiang)

This study discovers teachers' problems in designing Summative Test at Senior High School Kepahiang. This descriptive quantitative research, discusses the problems of teachers at Senior High School Kepahiang. In obtaining data, 62 teachers were given a questionnaire to determine teachers' problems in designing english Summative test. After getting the data, the researcher analyzed the teachers' responses by calculating the frequency and percentage. This study shows that the majority of teachers already had a good level of understanding of the procedures of designing a summative test. The teachers' levels of understanding were portrayed on their responses to the questionnaire items classified based on several themes or indicators taken from the constructed theories, including 1) characteristics of a good test, 2) planning a test, and 3) designing a test. In addition, the study also shows that majority the teachers had some problems in several aspects, namely 1) class and student communication frequency and 2) feedback from students to teachers to understand the students' level of understanding.

Keyterms: Teachers' problems, Designing, and Summative test.

MOTTO AND DEDICATION

“Your not a tough fighter. But no matter how many times you fail, you always find a way to rise up and fight”

Every miracle has its own way to solve, this research will never accomplish without many people helps. My best thanks to my advisor, Mrs. Desfitranita, M.Pd and my co-advisor Ms. Henny Septia Utami, M.Pd who had given so many contributions for my thesis.

This Thesis Dedicates To:

- My wonderful and beloved family, the greatest man ever my father Mr. **Sundar Mawan**, the most patient and beautiful woman ever, my mother Mrs. **Haya Lela Wati**.
- My beloved brother, **Yudahari Sandi**, **Prayoga Syahputra**, and **Yusrialtaf Musyaffa** who always become my partner and my sister **Ema Romantika** who always support me.
- My great advisor Mrs. **Defitranita**, M.Pd who always give me motivation, suggestion, guidance in finishing my study at IAIN Curup.
- Myco-advisor Ms. **Henny Septia Utami**, M.Pd who gave the researcher guidance, support and suggestion in finishing this thesis and also in finishing my study.
- My beloved lecturers of English Tadris Study Programe in IAIN Curup for support, suggestion and advices.
- My grandma no **Siti Muna (Siti)**, my grandfa nang **Zainul Abidin (Inong)** and nang **Sanir**. Thank u love for every single thing you give tome, support and du'a.
- My cousin **Lidia Chintya Dewi** and **Sulistri** thank you so much for your support and kindness.
- My beloved lovely friend **Ade Elcha Putri Siska** thank you so much dear for your support, kindness, and love for every single thing you give tome.
- My Great thanks to my beloved Family.
- My sister **Ema Romantika**, thank you for all of thing you give, your support, motivation, kindness, and your money for my life. You are my ATM when I'm stuck, May be (I'm crazy without you) thank you for giving me meaning to complement each other, you always take care of me when I'm seriously ill without asking for anything in return. No one word can describe you then "You are the best woman and sister"

- My best Support System, Muhammad Farid Ananda. Thank you for all advice and motivation.
- All of my friends in English Tadris Study Program Academic Years 2015 especially A class. I can't mention one by one. Thankyou so much for your help and I'm sorry if I have ever made something that make you offended, disappointed, and broke heart when we together!
- All of students' English Study Program IAIN Curup.
- My Almamater IAIN Curup.
- All of the people around me who gave me the golden precept that I could not be mentioned one by one.*

LIST OF CONTENTS

TITLE PAGE	i
SUPERVISORS' APPROVA	ii
THE STATEMENT OF OWNERSHIP	iii
PREFACE	iv
ACKNOWLEDGEMENT	v
ABSTRACT	vii
MOTTO AND DEDICATION	viii
LIST OF CONTENTS	xi
LIST OF TABLES	xiii
 CHAPTER I INTRODUCTION	
A. Background of the research	1
B. Research question	4
C. Objective of the research	4
D. Operational definition	4
E. Delimitation of the research	5
F. Significance of the research	6
 CHAPTER II LITERATURE REVIEW	
A. Review of the related theory	7
1. Test in teaching	7
2. Summative test	16
3. Designing summative test	27
4. Problems in designing summative test	31
5. Review of related findings	31
 CHAPTER III RESEARCH METHODOLOGY	
A. Kind of research	35
B. Population and sample	36
C. Technique of collecting data	38
D. Research of collecting data	40
E. The technique of analysis	49
 CHAPTER IV FINDING AND DISCUSSION	
A. Finding	54
B. Discussion	65
 CHAPTER V CONCLUSION AND SUGGESTION	
A. Conclusion	68
B. Suggestion	69
 REFERENCE	
 APPENDIX	

LIST OF TABLE

Table 2.1 The Rules in Designing Multiple Choice Items

Table 3.1	The Sample of the Research.....	37
Table 3.2	Instrument of Questionnaire 1.....	45
Tabel 3.3	Instrument of Questionnaire 2.....	48
Table 4.1	Teachers' Understanding of the Procedures of Designing a Summative Test.....	53
Table 4.2	Teachers' Problems in Designing a Summative Test.....	54

CHAPTER 1

INTRODUCTION

A Background of the Research

The teaching and learning process has a target that students can master english skills properly. Students' achievement become feedback for teachers and students. To measure student's achievement for previously given material, a learning evaluation is needed. Therefore, the teaching learning process and the evaluation must not be separated from each other.

Evaluation concerns all of the activities to improve quality, productivity, or performance in an institution during the teaching learning process. Especially in the education field, evaluation is needed to find information that focus on students achievement reached by a group or class. The students who had a satisfying or unsatisfying result will be motivated to increase their skill. However, some students who get good grades tend to be lazy to study and feel satisfied with their accomplishment and the students who fail on his assessment will lose their motivation to study. The information through the evaluation will be used by teachers to anticipate and prevent this learning process failing, so they can develop their teaching methods to increase their performance and student's achievement. They also can consider whether the teaching and learning process is successful or unsuccessful.

There are techniques can be applied to find information throught that evaluation. One of them is an achievement test. Huges demonstated that achievement test have two types, the first is final achievement test and the second

is progress achievement test.¹ That statement presents an agreement that measuring student ability use two kinds of achievement test. The first is final achievement test which is comprehended by summative test. The last test is progress achievement test or formative test. Two kinds of this test must be pleasantly constructed by a teacher who wants to measure their students ability.

Designing a good test is exhausting and need a guide. It needs a hard work for the teacher. Brown sets out several stages of test construction which consist of determining test objectives, drawing up test specifications, devising test task, scoring, grading, and giving feedback. To produce a better one, a teacher must follow the available syllabus and deal with many references related to the rules on how the test items should be made. As a consequence, the teacher is not allowed to make a test based on his desires without referring to the syllabus. Testers need to understand the test takers, their previous experience and background, and their abilities.²

A test made by teachers in the classroom should be in line with the syllabus, and the content of the test must measure what is intended to measure, in this case is the language skill taught by teachers. Is discussing the validity, the validity of a test is the extent to which a test measures what is intended to measure. Without validity there can be no confidence in the inferences and conclusions made from the results. It can be understood that validity hold enough

¹ Athur Hughes, *Testing for Language Teacher*, (Great Britain: Cambridge University Press., 2003), p.13

² H. Douglas Brown, *Language Assessment Principles and Classroom Practices*, (Longman : London, 2004), p. 3

important roles, it is seen from the content of its validity that is the test made by examiner must be able measure a competence.

Content validity is important to analyze the test. When the test which are given to the students is not valid, there are some effects. Firstly, the test can not measure the students achievement sufficiently. Secondly, the students are stiff to find the answers of the test because its material is not boned yet. The last, there is no concurrence between test item and material learnt which cause the students hardly to reach high score in the test.

Based on the pre-observations the writer had done at Senior High School Kepahiang, the writer found some interesting cases. First, there are some summative test sheets identical to the tasks in the printed book that students have studied, some test items or summative tests sheets were downloaded directly from the internet, and some test items served blurry pictures. Second, from the summative test assesment, there were several classes in which nearly all of his students failed the exam and required additional exams. Third, through open interviews with some teachers, the answers are they have constructed the test matter right, whether the tests are found on the internet or in printed books, the materials have already been studied by students. Therefore, the writer is interested to make a research on the research title **“Teachers’ Problems in Designing Summative Test at Senior High School Kepahiang”**.

B Research Questions

Based on the background of the research mentioned above, the problem of this research are :

1. How is teachers' understanding of the procedures of designing a summative test?
2. What are the problems faced by teachers in designing a summative test?

C The Objective of Research

The objectives of this research are to investigate :

1. Teachers' understanding of the procedures of designing a summative test.
2. Teachers' problems in designing a summative test.

D Delimitation of the Research

The researcher limits the study discussing as follow :

1. The research focus on the understanding and problems of teachers in designing summative test.
2. The test being analyzed is the English summative test for senior students of Senior High School which are placed in Kepahiang.
3. The research centered on reading and writing skills only because the summative test agree with particulars of multiple choice and essay.

E Definition of Key Terms

1. Teachers' problems

Teachers' problem in the study of science studies are often defined as the gap between expectation (aspired) and reality (generated).³ In this research, teachers' problems were defined as some challenges faced by the teachers or their problems in designing summative test items to assess student's english achievement at Senior High School Kepahiang.

2. Designing Summative test

Design was defined as an action to work out the details of something. The essential meaning of design is to plan and make decisions about something that is being built or created.⁴ Test is a method of measuring a person's ability, knowledge, or performance in a given domain.⁵ The test method used have to be structured by several techniques that provide an accurate measure or provide an explicit method. The test measure ability, knowledge, or performance in general or focus on specific competencies. Test are a subset of assessment. Assessment is an ongoing process that includes a much wider dominan. Whenever students offers a comment or answer the questions, these include an assessment that must be remarked by the teachers.

³ As Homby, *Problems In The Delivery Of The School-Based Psychoeducational Services To The Asian Immigrant Children*, Journal Children In Contemporary Society, 15 (3), (1983), p. 89

⁴ Merriam-Webster.Com, "Definition of Design" at <https://www.merriam-webster.com/dictionary/design> Retrieved on Monday, 6 September 2021

⁵ H. Douglas Brown, *Language Assessment Principles and Classroom Practices*, p. 3

Summative test or final achievement test are those administered at the end of a course of study.⁶ In sum, designing a summative test is an action by teachers to plan and make decisions about final achievement test.

F The Significance of the Research

The result of this study is expected to give a description for the readers about teachers' problems in designing summative test at Senior High School. It also can be used as an input for the readers; especially for the English teachers, the headmaster, and all people who are involved and responsible in developing quality of education. In other word, it is useful for all people to know the characteristics of a good test and for the researchers as the basic for conducting further research.

⁶ Athur Hughes, *Testing for Language Teacher*, p. 8

CHAPTER II

REVIEW OF RELATED LITERATURE

A. Test in Teaching

1. The Concept of the Test

Test is a formal, systematic procedure used to gather information about students' achievement or other cognitive skills.⁷ In addition, testing is primarily about establishing ways of making decisions that are (hopefully) not random, and seen as fair by the population.⁸ The procedures are systematic that measure person's ability or knowledge in agiven domain.

Generally, testing has several purpose, such as:

- a. To measure language proficiency
- b. To discover how successful students have been in achieving the objectives of a course of study.
- c. To diagnose students' strengths and weaknesses, to identify what they know and what they do not know.
- d. To assist placement of students by identifying the stage or part of a teaching program most appropriate to their ability.⁹

2. The Important of the Test

As one of the important part in teaching and learning process, the main important of test is to measure students' achievement toward the material given and inform the teacher how far the material has been understood by the

⁷ Peter W. Airasian & Michael K. Russel, *Classroom Assesment; Concepts and Applications*, (New York: McGraw-Hills companies, 2012), p.11

⁸ Glenn Fulcher, *Practical Language Testing*, (Britain: Hodder Education, An Hachette UK Company, 2010), p.4.

⁹ Athur Hughes, *Testing for Language Teacher*, Op. Cit., p.8

students, as describe by Bostwick and Gakuen state that test in an assessment can be used to improve instruction and help students take control of their own learning. That is more likely to be accomplished when assessment is authentic and tied to the instructional goals of the program.¹⁰ According to the explanation above, the test generally become a guidance for students to anlyze their strength and weaknes as a basis for improving their ability and to learn more about the material which still have the lower score by giving the real information to the students based on their performance. The decision will be taken by teachers or schools influence the prospect and opportunities of students to be success and pass the exams. In the other hand, the most significance of the test are to provide feedback and helping students to learn.

It can be concluded that the test is done as an effort to evaluate students' progress and improvement by giving the real information to the students based on their performance. In assessing students ability, the teacher does not work freely but there are some points that must be considered in order to get the fair result of evaluation. It includes the assessment types in assessing students and the teachers' technique in assessing students that relates to how the teacher gives the score to students' performance based on the criteria. Through these three aspects, the effective assessment is hoped to be reach.

¹⁰ RM Bostwick and Gakuen K, *Evaluating young EFL Learners : Problems and Solutions*, In Brown, J. D. And Yamashita, S. O. (eds), *JALT Allied Materials Language Testing in Japan*, Tokyo: The japan Association for Language Teaching. p.57

3. Kinds of Test

Arthur Huges in “*Testing for Language Teacher*” state four types of test, proficiency test, achievement test, diagnostic test, and placement test.¹¹

a. Proficiency Test

According to J.B. Heaton that the proficiency test is concerned simply with measuring a student’s control of the language in the light of what he or she will be expected to do with it in the future performance of a particular task.¹² Whereas James Dean Brown uttered that a proficiency test assesst the general knowledge or skills commonly required or prerequisite to entry into (or exemption from) a group of similar institution.¹³

Proficiency test or aptitude test is a type of tests designed to measure person’s ability to use language, regardless of his education in that language. Therefore, the content of this test, is not based on the content or objectives of the language courses taken by the participant, but based on a specification of the completed courses. Applicants must have language skill to be recognized as a professional.

Proficiency test is widely used for grading or selection, and its relative value lies in the ability to assign students to required subjects according to their abilities.

¹¹ Ibid., p.12-16

¹² J. B. Heaton, *Writing English Language Tests*, (USA: Longman inc., 1988), p.172-173

¹³ James Dean Brown, *Testing in Language Programs*, (New Jersey: Prentice Hall Regents, 1996), p.10

b. Achievement test

As the name suggested, the test is used to determine the extent to which individual student, the student group, or the course itself achieves its goals. H. Douglas Brown stated that an achievement test is related directly to classroom lessons, units, or even a total curriculum. This test is limited to particular material covered in a curriculum within a particular time frame, and is offered after a course has covered the objectives in question.¹⁴

According to Mehres and Lehmann stated on the book *A guide to language testing: Development, Evaluation and Research*, achievement test may be used for program evaluation as well as for certification of learned competence. It follows that such tests normally come after a program of instruction and that the components or items of the tests are drawn from the content of instruction directly.¹⁵

It can be concluded that achievement test is used to measure the learning level in a specific content area, and it is usually consistent with the clearly stated course objectives. Many times, when students know that they will take the weekend test or the semester achievement test, the result is an increase in study time, approaching the hour test.

¹⁴ H. Douglas Brown, *Teaching by Principles; An Interactive Approach to Language Pedagogy*, (New York: A Pearson Education Company, 2015), 4th. Ed., p. 500

¹⁵ Grant Henning, *A Guide to Language Testing: Development, Evaluation and Research*, (China: Heinle & Heinle Publisher, 2001), p.6

According to Arthur Huges, there are two types of achievement test:¹⁶

a. Summative Test (Final Achievement Test)

Summative assessments are efforts to use information about students or programs after a set of instructional segments has occurred. Their purpose is to summarize how well a particular student, group of students, or teacher performed on a set of learning standards or objectives. Information obtained from summative assessments is used by teachers to determine grades and to explain reports sent to students and their parents.¹⁷

In summative testing, it is expected that test scores to carry generalizable meaning; that is, the score can be interpreted to mean something beyond the context in which the learner is tested.¹⁸

It is concluded that, summative test is conducted at the end of the course and can be formulated and implemented by the staff of the ministry of education, the official examination committee or the educational institution. The purpose of this test is to understand how successful students learned the previous material during a long study period.

¹⁶ Arthur Huges, *Testing for Language Teacher*, Op. Cit., p.13

¹⁷ Richard I. Arends, *Learning to Teach*, (New York: McGraw Hills, 2012), 10th. Ed., p.221

¹⁸ Glenn Fulcher, *Practical Language Testing*, Op. Cit., p.20

b. Formative Test (Progress Achievement Test)

This is measure of the progress made in completing the final performance test, with an expected increase of points to show that progress has been made.

Formative tests take place while interacting with students and focused on making quick and specific decisions about what to do next in order to help students learn. They all rely on information collected through either structured formal activities or informal observations made during the process of instruction.¹⁹

Formative tests are typically designed to measure the extent to which students have mastered the learning outcomes of a rather limited segment of instruction, such as a unit or a textbook chapter. These tests are similar to the quizzes and unit tests that teachers have traditionally used, but they place greater emphasis on (1) measuring all of the intended outcomes of the unit of instruction, and (2) using the results to improve learning (rather than to assign grades).²⁰

The formative test results provide the information about the students' mastery of a subject. The purpose is to determine the success and failure of students for learning and learning adjustments. The formative test also determines whether the student has not mastered the learning tasks he taught, it can guide him to correct the learning errors.

¹⁹ Peter W. Airasian & Michael K. Russel, *Classroom Assesment; Concepts and Applications*, Op. Cit., p.99-100

²⁰ Arthur Hughes, 2003, *Testing for Language Teacher*, Op. Cit., p.15

c. Diagnostic Test

Diagnostic test is designed to determine the degree to which the specific instructional objectives of the course have been accomplished.²¹ The diagnostic test is widely used; few tests are constructed solely as diagnostic tests. Note that diagnostic testing is frequently carried out of groups of students rather for individuals.²²

Therefore, diagnostic tests are designed to diagnose specific aspect of the language and can be used to verify that students are learning specific element of the course. For example: it can be used at the end of a chapter in the text book or after completing a specific course.

d. Placement Test

The placement test provides an invaluable aid for placing each student at the most beneficial position in the instructional sequence.²³

The purpose of placement test is to place a student into an appropriate level or section of a language curriculum or school. A placement test typically includes a sampling of material to be covered in the curriculum (that is, it has content validity), and it thereby provides an indication of the point at which the student will find a level or class to be neither too easy nor too difficult, but appropriately challenging.²⁴

Therefore, the placement test is designed to provide information to help students reach the level or part of the teaching plan that best suits

²¹ James Dean Brown, *Testing in Language Programs*, Op. Cit., p.15

²² J.B. Heaton, *Writing English Language Tests*, Op. Cit., p.173

²³ Norman E. Gronlund, *Constructing Achievement Tests*, (USA: Prentice Hall Inc., 1977,), p.3

²⁴ H. Douglas Brown, Loc. Cit.

their abilities. The classroom teacher may know that students are ready to start learning and incorporate it as part of classroom activities under proper guidance.

4. Characteristics of a Good Test

A test design must identify different tests based on the purpose, time, topic, and the most specific characteristics of a good test. If a test has a certain score on the four characteristics that you are looking for, then it can be called a Good test. Effectiveness, reliability, predictability and objectivity, teachers can motivate students to improve their learning.

1) Practical

An effective test is practical. This means that it is not excessively expensive, stays within appropriate time constraints, is relatively easy to administer, and has a scoring or evaluation procedure that is specific and time-efficient.²⁵

2) Reliability

A reliable test is consistent and dependable. It means if the test is given to the same student or matched students on two different occasions, the test should yield similar results.²⁶

In the other hand, reliability test is affected by a number of factors, chief among them being the adequacy of the sampling of tasks. Poor students' motivation can be attributed to weaknesses in the test or the

²⁵ Douglas H. Brown, *Language Assessment Principles and Classroom Practices*, Op. Cit., p. 19

²⁶ Ibid., p. 20

testing procedure, and sometimes it is caused by factors beyond either the test writer's or examiner's control.²⁷

In sum, Reliability refers to the stability of test result. If the measurements are consistent, the test cannot measure well. For example, in order to be safe on the meter, we need to ensure that the results of about are the same.

3) Validity

Validity is the most complex criterion of an effective test and arguably the most important principle.²⁸ If the test results are based on a reliable analysis of the skills we want to measure, and if there is enough evidence that the test score are closely related to the actual skills in the tested skill area, then we can feel quite safe. As long as the test valids to our purpose.

4) Authenticity

Authenticity is the degree of correspondence of the characteristics of a given language test task to the features of a target language task, and then suggest an agenda for identifying those target language tasks and for transforming them into valid test items.²⁹

²⁷ David P. Harris, *Testing English as a Second Language*, (New York: McGraw Hill Book Company, 1969), p. 14

²⁸ Douglas H. Brown, *Language Assessment Principles and Classroom Practices*, Op. Cit., p. 22

²⁹Ibid., p. 28

5) Washback

Washback generally refers to the effects the tests have on instruction in terms of how students prepare for the test.³⁰

In brief summary, as an evaluator or teacher of development evaluation, there should be something available in the evaluation. First, when choosing a test, the teacher may think, "What exactly does the test measure? How effective is the test measurement? Second, whether the test is practical or not. Third, the topic selection should be contextual and meaningful. The latter is the impact of examinations on teaching and learning.

B. Summative Test

1. Concept of Summative Test

Summative test takes place at the end of a large chunk of learning, with the results being primarily for the teacher's or school's use. Alberta said Summative test is usually conducted at specific times, after students have had opportunities to practice, at the end of a unit or semester or at the end of reporting period. Summative test typically are given to students at the end of a set point during or at the end of the semester to assess what has been learned and how well it was learned.³¹ Teachers or schools can use this test to

³⁰ Ibid.

³¹ Alberta. *Formative and Summative Assessment*. (Chicago: Northern Illinois University of Chicago Faculty Development and Instructional Design Centre, 2008). Unpublished. Retrieved from [Www. HOI science.com](http://www.HOIscience.com), on August 20th 2018

identify of curriculum and instruction with improvement affecting the term's student. Summative test can be done through paper and pencil tests, unit tests, grading of student assignments, presentation and projects, grading of student portfolios.

The types of summative test are:³²

- a. Examinations
- b. Final Examinations
- c. Term papers
- d. Project
- e. Portfolios
- f. Performance
- g. Students evaluation of the course

The summative test records learning at the end of an instructional segment, discourage the students and the effect is weak and fleeting. Moreover, summative tests are used to evaluate student learning skill acquisition and academic achievement at the conclusion of a defined instructional period. This test as mean of the measuring student's achieving after having training process. It is the formal testing of what has been learned in order to produce marks or grades which may be used for reports of various types and given periodically to determine at a particular point in time.

³² *Formative and Summative Assessment*. (Chicago: Northern Illinois University of Chicago Faculty Development and Instructional Design Centre, 2008). Unpublished. Retrived from Www. HOI science.com, on August 20th 2018

2. Subjective Test

Subjective tests mostly used during the intuitive era and later on the objective ones have been often used now since the scientific and communicative era. Tinambunan stated that subjective test items present a less structured task than objective type items, and consequently it is more difficult to control the nature of the student's response.³³

Subjective test is generally in the form of essay question or rather long supply-type item. In essay test, the tester must think carefully of what to say and then express ideas as well as possible. The subjective judgment of scores enters into the scoring, and thus, the scores differs from one scores to another and from one time to another the same scorer.

3. Objective Test

The objective test includes a variety of forms of test tasks having in common the characteristic that the correct answer, usually only one, is determined when the test item is written. Thorndike and Hagen Stated the word "objective" in objective test refers only to the scoring of the answers; the choice of content and coverage of an objective test is probably as subjective as the choice of content and coverage of an essay test, and for some types of items there is subjective judgment involved in the original decision as to what is the correct answer.³⁴

Karmel stated the objective test is so called objective because the scoring procedure is determined when the test is written. That is, the correct answer, usually only one, is, completing stated before testing. Thus the grader can be completely objective about the answer.³⁵ The objective test is a structured examination. That is, each examinee is presented with exactly the same

³³ Wilmar Tinambunan, *Evaluasi Hasil Belajar*, (Jakarta: Departemen Pendidikan dan Kebudayaan, 1988) p.34

³⁴ RL. Thorndike and Hagen E., *Measurement and Evaluation in Psychology and Education*, 2nd Ed, (USA: G and C Meriam Co, 1962), p. 47

³⁵ L. J. Karmel and Karmel M. O., *Measurement and Evaluation in Teaching*. 2nd Ed, (USA: Macmillan Publishing Co, 1978) p. 420-421

problem. The objective, on the other hand, being completely structured, must be answered in a prescribed manner. The student is not called upon to organize his response as he is in the essay format. The objective test requires the student to recognize, not to recall, the correct answer. This is because most objective tests present given alternatives (with the exception of the completion item), one of which is the correct response.

4. Types of Objective Items

According to Karmel, there are four types of objective items, there are:³⁶

a. True-False Items

The true-false item has been very popular with teachers, probably because it is easy to construct and requires little time. The following statements are representative of the major drawbacks of the true-false item:

- a) The true-false item tends to be greatly influenced by guessing.
- b) It is almost impossible to make statements either absolutely true or absolutely false.
- c) True-false tests foster poor test-taking habits. Students are clever and will second-guess the teacher who employs the true-false item and discern pattern.

b. Completion Items

Completion items require the student to fill in a blank that completes the sentence or answer a specific question. The completion item is related to the essay item and serves as a bridge between the objective and essay test. On the one hand, it is objective, in the sense that a prearranged answer can be chosen before testing; on the other hand, it is

³⁶ L. J. Karmel and Karmel M. O., Op. Cit., p. 422-423

related to the essay test because the student must produce the correct answer rather than recognize it. The completion item is especially useful for appraising your student's knowledge of facts, such as names and dates.

c. Matching Items

The matching item's major advantage is that it condenses a great deal of material into a limited amount of space. The matching item is simply a modification of the multiple-choice form. Instead of the possible responses being listed underneath each individual stem, a series of stems, called premises, is listed in one column and the responses are listed in another column.

d. Multiple-Choice Items

The multiple-choice format is one of the most popular and effective of all the objective tests. It consists of two parts: (1) the stem, which states the problem, and (2) a list of options, one of which is to be selected as the answer. The stem may be stated as a question or as an incomplete statement. The multiple-choice item can be used to appraise almost any educational objective with the exception, of course, of student organization and ability to produce answers.

C. Designing Summative Test

1. Planning the Tests

Design and conduct summative test is a part of instructional design. Dick and Carey stated ten phases of instructional design which ends with the design of summative test.³⁷

³⁷ Walter Dick, Lou Carey, James O. Carey, *The Systematic Design of Instruction*, (USA: Library of Congress Cataloging, 2015), p.6-8

- a. Identify Instructional Goal
- b. Conduct Instructional Analysis
- c. Analyze Learners and Context
- d. Write Performance Objectives
- e. Develop Assessment Instruments
- f. Develop Instructional Strategy
- g. Develop and Select Instructional Materials
- h. Design and Conduct Formative Evaluation of Instruction
- i. Revise Instruction
- j. Design and Conduct Summative Evaluation

Summative test occurs only after the instruction has been formatively evaluated and sufficiently revised. The summative test that made by teachers should pay attention to the formative test and other part of instructional design. So, before design a test, the teachers must determine assessment goals, competence standars, and basic competence that will be measured. Depdiknas stated some instructions before design a test, that are:³⁸

- a. Identify the assessment goals. The porpuses of assessment is crucial because each assessment has different purposes, for example the purposes of learning achievement tests, diagnostics test, placement test, or proficiency test.
- b. Focus on Standard Competence (SK) and Basic Competence (KD). Competence standard is a primarily target should be measured through each basic scompetences.
- c. Identify the test instrument. The test materials should consider the urgency (the materials should be mastered by the learners),

³⁸ Departemen Pendidikan Nasional, *Panduan Penulisan Butir Soal*, (Jakarta: Depsiknas, 2008), p. 7-8

continuity (advanced materials), relevance (have beneficial to other subject), and have beneficial in learners' daily. The teachers also should pay attention to the form of the test, for example oral tests or written tests.

- d. Design the examination grid with its suspension guidelines.

In addition, in constructing a summative test to fit a set of specification, the test maker may choose from variety of item types. Some of the test items are referred to as objective items, because they can be scored objectively. That is, equally competent scorers can score them independently and obtain the same result. They also include the following selection-type items: multiple-choice, true-false, and matching. They also include the supply-type items that are limited to short answers (several words or less), even though such items are not completely objectives. The other supply- type item, the essay question, is subjective. That is, the subjective judgment of tdhe scorer enters into scoring, and thus, the scores differ from one scorer to another for the same scorer.

2. Designing Objective Test

Designing a summative test, almost teachers especially in Senior High School agreed to use multiple-choice items and essay as the instrument to assess students' achievement. Some of them also designing matching items as a modification of multiple-choice items.

a. Designing Multiple-Choice Items

The multiple-choice item is generally recognized as the most widely applicable and useful type of summative test item. It can more

effectively measure many of the simple learning outcomes measured by the short-item or completion, the true false item and the matching item. It can measure a variety of the more complex learning outcomes in the knowledge, understanding and application areas.

Gronlund stated a multiple-choice item consists of a problem and a list of suggested solutions. The problem may be stated in the form of a direct question or an incomplete statement and is called the stem of the item. The list of suggested solutions may include words, numbers, symbols, or phrases and are called alternatives. The pupil is typically requested to read the stem and the list of alternatives and to select the one correct, or best, alternative. The correct alternative in each item called merely answer, while the remaining alternatives are called distracters.³⁹

According to Brown, there are four criteria in constructing multiple choice test. The four criteria are:⁴⁰

- a) Design each item to measure a specific objective.
- b) State both stem and options as simply and directly as possible.
- c) Make certain that intended answer is clearly the only one correct answer.
- d) State the stem of the item in positive form, wherever possible.

³⁹ N. E. Gronlund, *Measurement and Evaluation in Teaching, Fourth Edition*, (New York: MacMillan, 1981), p. 178

⁴⁰ H. D. Brown, *Language Assessment. Principles and Classroom Practice*, Op. Cit., p. 55-58

Bloom also suggest five criteria for constructing multiple choice test. The criteria as stated below:⁴¹

- a) Have all unintentional clues been avoided?
- b) Are all of the distracters plausible?
- c) Has needless redundancy been avoided in the options?
- d) Has the ordering of the options been carefully considered? Or are the correct answer randomly assigned?
- e) Have distracters like “none of the above,” A and B only”, etc. been avoided?

Design a multiple-choice items requires precision and skill. Based on the Ministry of Education, there are some rules in designing the multiple-choice items. They can been sen on the following table.⁴²

Table 1
The rules in designing Multiple Choice Items

The Material Aspects	a. The items are matched to the indicators
	b. Make sure there is only one anwer key for each item.
	c. The material is matched to the assessment goals.
	d. The material is appropriate tho the students' level.
	e. The distractors are available in the answers.

⁴¹ B. S. Bloom, *Taxonomy of Educational Objectives*, (New York: Longman, 1956), p. 48-50

⁴² Departemen Pendidikan Nasional, *Pedoman Umum pengembangan Penilaian*, (Jakarta: Depdiknas, 2003), p. 76-80

The Construction Aspects	f. State the stem of the item in simple, clear language.
	g. Present a single clearly formulated problem in the stem of the item.
	h. Avoid verbal clues that might enable students to select the correct answer or to eliminate an incorrect alternative.
	i. State the stem of the item in positive form, wherever possible.
	j. Emphasize negative wording wherever it is used in the stem of an item.
	k. Make all alternatives grammatically consistent with the stem of the item.
	l. Avoid using the alternative “all of the above,” and use “none of the above” with the extreme caution.
	m. Vary the relative length of the correct answer to eliminate length
	n. Sort the answer based on the numbers and times.
	o. Use the effective item format.
	p. Make certain each item is independent of the other items in the test.
Language Aspects	q. Present a communicative language.
	r. The statement present a common language.
	s. Present a clear statement to avoid misunderstanding to the item.
	t. The statement is not containing the offensive words.

b. Designing Matching Items

Some teachers prefer to use matching items, because it is a simply modification of the multiple-choice form. Gronlund satated some rules for constructing matcing items, that are: ⁴³

⁴³ Norman E. Gronlund, *Constructing Achievement Tests. Second Edition*, (USA: Prentice-Hall Inc., 1977), p. 57

- a) Include only homogeneous material in each matching item.
- b) Keep the list of items short and place the brief response on the right.
- c) Use a larger, or smaller, number of responses than premises, and permit the responses to be used more than once.
- d) Specify in the directions the basis for matching and indicate that each response may be used once, more than once, or not at all.

2. Designing Subjective Test

Subjective test is generally in the form of essay question. Constructing essays test requires precision of the test maker. The accuracy of the test requires students to organize their ideas in written form. Constructing essays test is divided into three rules, that are:⁴⁴

a. The materials

The materials of essays test must consist of:

- 1) The items are matched to the indicators
- 2) Every questions must be given the expected answer limit.
- 3) The items are matched to the assessment goals.
- 4) The material is appropriate tho the students' level.

b. The Constructions

In designing an essay test, the test maker should agree with these following rules.

⁴⁴ Departemen Pendidikan Nasional, *Panduan Penulisan Butir Soal*, p. 14

- a. The items use a question/command that demands the unravelled answers.
 - b. Present a clear statement for each item.
 - c. Each items should have its suspension guidelines.
 - d. Use the effective item format.
- b. The Languages

The test makers should pay attention to the language that they used as following rules.

- a. The items present communicative language.
- b. The items present a common language.
- c. Avoid the ambiguous sentence.
- d. The statement is not containing the offensive words.

D. Problems in Designing Summative Test

1. Teachers problems on Assessment

The term problem / problematic originates from Latin, namely “problematic” Where is in the Indonesian dictionary, the problem means things that cannot be solved; which causes difficulties to solve it.⁴⁵ As for the problem itself “is an obstacle or problem that must be solved in other words the problem is the gap between reality and something that is expected to be good, in order to achieve maximum results,” come from internal or external factors.

⁴⁵ Kahneman, D., & Miller, D. Norm Theory: Comparing Reality to Its alternatives. *Psychological review*, 93, (1986), p. 136

As Homby argues that teachers' problems in the study of science studies are often defined as the gap between expectations (aspired) and reality (generated). Thus an effort is needed to aim at something as expected.⁴⁶ Learning success is largely determined by how far the teacher is able to minimize or resolve existing problems. The fewer problems the greater the chance of student learning success, and vice versa, in addressing or perceiving the learning process. This problem arises from the teacher's perspective on the teacher's role and the meaning of learning. Third, social problems, namely problems related to the relationship and communication between teachers and other elements outside the teacher, such as the lack of harmony between teachers and students, between school leaders and students, even among fellow students.⁴⁷ The disharmony between teachers and students can be caused in addition to cultural factors which can also be caused by patterns or system of leadership that lack democracy or pay little attention to humanitarian issues.

Fu Ren suggested some ways to build a harmonious relationship between teachers and students, that are:⁴⁸

- a) Class and students communicative frequently, through the communication with the students to show the school practices and requirements, close the distance to the students, communicate feelings.

⁴⁶ As Homby, *Problems In The Delivery Of The School-Based Psychoeducational Services To The Asian Immigrant Children*, p. 89

⁴⁷ Ibid., P 90

⁴⁸ Fu Ren, *The Construction of a Harmonious Relationship Between Teachers and Students to Improve the Actual Effect of Moral Education*, Journal of Social Science, Vol. 11, No. 5, p. 13-17

- b) Class management democratization, teachers give full play to students' enthusiasm and creativity, realize the harmonious unification of class, and democratic class.

2. Educational Partnership in enhancing student assessment

The cooperation between different participators at the local level and the transparency of the evaluation process are also important. Involving students in the assessment process is an effective way to increase student confidence, because as confidence increases, so does belief in learning.

If teacher fail to verify their observations or grades, students may not be able to obtain evaluation assistance. Russell and Austin in their research stated that many teachers will independently develop their own assessment methods without considering the assessments methods of their colleagues.⁴⁹ Teachers should share and evaluate the assessment strategy with their colleagues. They can share information and develop their understanding of the assessment process between and within schools, enabling them to redefine their own teaching practices, student learning, and understanding of subject goals.⁵⁰

3. Feedback in Supporting Learning

Feedback is an important part of the learning cycle and lies at the core of the assessment. Versatile feedback is an element in ensuring transparent and therefore fair assessment processes. Feedback should follow a three way

⁴⁹ Joshua A. Russel & James R. Austin, *Assessment Practices of Secondary Music Teachers*, Journal of Research in Music Education, Vol. 58 No. 1, p. 37-38

⁵⁰ Wyne Harlen, *Teachers' Summative Practices and Assessment for Learning-Tension and Synergies*, The Curriculum Journal, Vol. 16 No. 2, 2005, p.214

path: from students to teachers so that the teacher can understand the students' level of understanding; from teacher to students, whereby the teacher responds to a challenge or extends the student's ideas and from student to student, in as much as students can help and be helped by mutual dialogue.⁵¹

The feedback provided by teachers to students seem to have a social and managerial role, usually at the expense of teaching. Collecting grades to complete the record takes priority over analyzing student performance to determine learning needs. Teachers do not know the grades of their students' past teacher.

4. Components of the Change Environment that Facilitates of Assessment Practices

There are four Factors that affect teachers' adoption of Assessment. First, personal factors related to teachers' personal and professional values and can change from one teacher to another. The factors as following the teachers' beliefs and values, teachers' knowledge or pedagogical content knowledge, teacher's attitudes, teachers' orientations, teachers' perceptions, and teachers' understanding of assessment. Second, contextual factors, are directly related with the teaching context that includes schools' environment and realities. The factors affect teachers' decisions about assessment. The policy and context of school, the support from the institute, students' attitude, and social cultural also following the contextual factors of teachers made-test. Third,

⁵¹ Paivi Atjonen, *Teacher's View of Their Assessment Practice*, The Curriculum Journal, Vol. 25 No. 2, 2014, p. 243

external factors, also affect teachers' classroom practices and decisions about assessment. There are state and local educational policies, high stakes and accountability assessment, and curriculum developers. The last, resource related factors, which are related to resources such as information, material, funding and time, affect the teachers' assessment adoption.⁵²

E. Review of Related Finding

The researcher finds other researchers about variable that have related to the object of the present reserach that the researcher did. But, the researcher could not access all of those studies completely. There are several previous reserach can be summarizing as the following:

The first reasearch that had be done by Desri Susiyanti on her thesis "Theachers' difficulties in Theaching English as a Foreign Language at SMK Muhammadiyah 3 Makassar".⁵³ She found many difficulties faced by the teachers such as lack of student's basic knowledge in English lesson, a lot of student in the classroom, the students' low concentration, the student's absent in learning process, boredom, and the facilities in the scholl was low.

The second finding is Isabell Sengkaton, "An Analysis of Theacher-Made Summative Test of english of SMA Negeri 3 Palembang".⁵⁴ He found the teacher-made summative test of Englishof the eleventh-grade students' of SMA Negeri 3

⁵² Kemal Isci, *Internal and External Factors affecting Teachers Adoption of Formative Assessment to Support learning*, International Journal of Social, Educational, Economic, Business, and Industrial Engineering, Vol. 10 No. 8, 2016, p. 4-8

⁵³ Desri Susiyanti, 2019, *Theachers' Difficulties in Theaching English as a Foreign Language at SMK Muhammadiyah 3 Makassar*

⁵⁴ Isabell Sengkaton, 2020, *An Analysis Of Teacher-Made Summative Test of English of SMA Negeri 3 Palembang*

Palembang was not acceptable to be used as a tool to evaluate the students' comprehensions because many of the items did not fulfill all aspects of a good test in terms of quality, index difficulty, discrimination power and effectiveness of distractors.

The third finding is Nurhalimah & Fahriany, in their journal "Determining the Quality of English Teacher-Made Test: How Excellent is Excellent?".⁵⁵ They found indicate that the English mid-term test has 24 acceptable items (80%) from the quality excellent, good, and satisfactory. Then, three items (10%) have poor quality, and three items (10%) have very poor quality, or in the negative value on discrimination index to the extent that the items are eliminated. It is proven by statistical data that they fail to distinguish between students who are knowledgeable and those students who are not on the base of how well they know the materials that have been tested.

Everlyn Olouch, the journal "Challenges Faced by Tutors in Setting of Examinations".⁵⁶ The findings are : during the moderation of summative examination it is noted that some tutors have challenges in setting examination. They require adequate support to acquire the necessary skills through workshop, seminars and mentoring by experienced colleagues. Providing proper orientation to tutors is mandatory for setting of effective examinations. Equipping them with the needed skills improves the quality of examinations in an institution and gives tutors confidence. Majority of tutors know the purpose of examinations which is

⁵⁵ Nurhalimah & Fahriany, *Determining the Quality of English Teacher-Made Test: How Excellent is Excellent?*, Journal of Linguistics, and Literature, Vol. 5, No. 1, 2019

⁵⁶ Everlyn Olouch, *Challenges Faced by Tutors in Setting of Examinations*, Journal of education, Vol. 5, No. 17, 2014

to test learners' mastery of course content but the problem is to do with doing it effectively. Most tutors also confirm that validity, testing what is supposed to be tested, is the major characteristics of a good examination. However, it was noted that the staff development workshops quality was eventually attained.

M. Aries Taufik, dkk, their research "Item analysis and Teachers' Factors in Designing a Test", accomplished at Junior High School in one of the regions in Riau.⁵⁷ They found the test was dominated by moderate items and those items are functioned well. However, most of the items cannot discriminate between high and low students. In addition, it is found that the first teachers' factor significantly affecting the effectiveness of the test items is training on test construction, the followed by teachers' experience in constructing a test. It is expected that the authority could select more experienced teachers to design the tests and frequently facilitate the teachers with training on test construction to enhance teachers' competence in designing tests.

Paivi Atjonen, the journal "Teachers' views of Their Assessment Practice",⁵⁸ found the majority of positive views concerned the use of different assessment methods, an interactive approach, encouraging feedback, and criteria clarification. Negative views dealt with improper assessment methods, level of stringency, badly implemented assessment, and weak ground of assessment. Three factors made assessment as difficult: interpretation of fairness, pupils with special needs, and pupil heterogeneity. On the other hand, versatile assessment methods,

⁵⁷ M. Aries Taufik, et.all., 2029, *Item Analysis and Teachers' Factors in Designing a Test*

⁵⁸ Paivi Atjonen, *Teacher's View of Their Assessment Practice*, The Curriculum Journal, Vol. 25 No. 2, 2014

curriculum advice, and pupils' good competencies made teachers' assessment work easy.

Based on the previous research, the researcher consulted that all of the research in advance was connected with this research. On the other hand, the first finding was about teachers' difficulties in teaching whereas this research is about teachers' problems in designing a summative test. It was being information which is necessary to the researcher for knowing some problems in other situation. The researcher had the difference aim but has similar scope. The next finding also had the relationship with this study because it was about summative test. Besides, majority of the researchs focus on the content analysis of summative tes. It has difference scope with this research that focus on the teachers who made the test. The last finding was quite relevant with this study where it told about teachers' views of their assessment that found some difficulties that faced by the teachers. It was quickly similar with this study but had the different place.

CHAPTER III

METHODOLOGY OF THE RESEARCH

A. Kind of the Research

Methodology of this research is technique and procedure tools which selected in execute research.⁵⁹ The method in this research is a descriptive quantitative research because the researcher only analyzed and interpreted the problems faced by English teachers in designing English Summative Test. Descriptive method explained the data which have correlation with fact, situation variable and phenomenon which happen when the research are being conducted. According to Arikunto, descriptive research studies were designed to obtain information concerning the current status of phenomena.⁶⁰ Descriptive method means a research focused in describing any situation or condition in population, systematically, factually, and accurately.⁶¹

The study are about any population, survey the society opinion, and other. In this research the researcher well describe fact of teachers' problems in designing English Summative test.

⁵⁹ T. Fatimah Djajasudrama, *Metode Linguistic, Ancangan Metode Penelitian dan kajian*, (Bandung: PT Eresco, 1993), p. 3

⁶⁰ Arikunto Suharsimi, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: Renika, 2009), p. 128

⁶¹ Nurul Zuriyah, *Metodology Sosial dan pendidikan, Teori Aplikasi*, (Jakarta: Bumi Aksara, 2006), p. 37

B. Population and Sample

1. Population

Population is a set (or collection) of elements, possessing one or more attribute of research.⁶² According to Donal Ary, Population is defined as all members of any well-defined class of people, events or objects.⁶³ The population of this research is all of the English Teachers at Senior High School Kepahiang who made the summative test and taught 10th, 11th, dan 12th Class. The population consist of eighteen school. Total teachers were 62 people.

2. Sample

Sampling is the process of selecting a number of individuals for a study in such a way that they represent the larger group from which they were selected. A sample is small proportion of a population selected for observation and analysis.⁶⁴ Total sampling was employed in this research, it means the research took all the population as the sample, it involved all English Teachers at Senior High School Kepahiang who made the summative test and taught 10th, 11th, dan 12th Class as sample of this research. According to Syukarman Syarnubi, total sampling is a kind of sampling technique by deciding all member of research population as the sample.⁶⁵ The number of sample was given in following table.

⁶² Nasution, *Metode Research*, (Jakarta: Bumi Aksara, 1998), p. 101

⁶³ Donal Ary, *Introduction to Research in Education*, (USA: Wadsworth Group A Division Of Thompson Learning Inc, 2010), 8th Ed., p.148

⁶⁴ Ibid, p. 148

⁶⁵ Syukarman Syarnubi, *Metodology penelitian Kuantitatif dan Kualitatif*, (Curup: LP2 STAIN Curup, 2011), p. 102

Table 2

The sample of the reseach

Schools	Number of English Teachers
SMAN 1 Bermani Ilir	2
SMAN 1 Kabawetan	4
SMAN 1 Kepahiang	4
SMAN 1 Merigi	4
SMAN 1 Muara Kemumu	3
SMAN 1 Tebat Karai	4
SMAN 1 Ujan Mas	3
SMAS Muhammadiyah Kepahiang	2
SMKN 1 Bermani Ilir	3
SMKN 1 Kepahiang	3
SMKN 1 Seberang Musi	3
SMKN 1 Ujan Mas	4
SMKN 2 Kepahiang	6
SMKN SPPN Bengkulu	3
MAN 1 Kepahiang	3
MAN 2 Kepahiang	4
MAS 01 Darussalam Kepahiang	4
MAS Al-Munawwaroh	3

Total	62
-------	----

The purpose of taking all member within the population as sample in this research because the researcher wanted to find out and describe about the problems faced by English Teachers in designing English Summative test, if all member of population was taken as sample certainly the researcher get the data completely. Moreover, based on the principle in deciding the sample size, larger sample is more represent the population and result more accurate. As Purwanto says, that the principle in deciding the sample size is the larger sample it is more represent the population, so it is more accurate. If the sample represent the population, the data which are taken from the sample more accurate to predict the conclusion.⁶⁶ Related to this theory in order to get more accurate data the researcher took all member of population as the sample.

C. Technique of Collecting Data

To answer two research questions above, the researcher collected the data by using questionnaires. It means questionnaire for the first research question and second research question.

1. Questionnaire

Questionnaire is one of the techniques for collecting data. Questionnaire is a group of written questions on paper or the others and given to the respondents to be answered without any intervention from the researcher.⁶⁷ In

⁶⁶ Purwanto, *Metodologi Penelitian Kuantitatif Untuk Psikologi Dan Pendidikan*, (Yogyakarta: Pustaka Pelajar, 2008), p. 226

⁶⁷ Danim, *Menjadi Peneliti Kualitatif*, (Bandung: CV. Pustaka Setia, 2002). P. 138

this research, questionnaire is very important to get more detailed information such as problems when design the test, the respondent would have difficulties to mention it one by one. Therefore, the researcher used questionnaire to get those. Furthermore, the respondent could have more time to fulfillit. Questionnaire can be open-ended or close-ended.⁶⁸ In addition, Creswell stated three kinds of questinnaire, open-ended, close ended, and open- and closed-ended questions.⁶⁹ Open-ended questionnaire is a list of questions that there is no standart answer to these questions, and data analysis is more complex.⁷⁰ It means the respondent need to write the answer in answer sheet. In the other hand, closed-ended questionnaire is the researcher poses a question and provides pre-set response options for the participat. Open- and closed-ended questions has all the advantages of open-ended and closed-ended questions.The technique is to ask a closed-ended question and then ask for additional responses in an open-ended questions.

The researcher used questionnaire to answer those two research questions. The format of questionnaire in this research is open- and closed-ended. It means the respondents answered the questions about based on the answers provided that has been prepared by researcher and the researcher also ask for the other response if the participant had different answers from the list provided.

⁶⁸ Sugiyono, *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*, (Bandung: Cv. Alfa Beta, 2010), p. 199

⁶⁹ Cresswell, 2012, p. 389-390

⁷⁰ Chaterine Dawson, *Practical Research Method*, (United Kingdom: How to Books, 2002), p. 138

In conclusion, the using of questionnaire here was open- and closed-ended with purpose to determine the teachers' problems in designing English Summative test and also the factors that cause the problems. The questionnaire was designed by the researcher with four options, the alternative options are:

- a) Always : 4
- b) Often : 3
- c) Seldom : 2
- d) Never : 1⁷¹

The researcher also add other optional if the respondents have their own answer.

Then, the data collection process by using questionnaire are as:

- a) The questionnaire is given to the respondents via google form;
- b) Collecting the questionnaire;
- c) The questionnaire tabulated and analyzed based on the percentage formula.

Based on the statement above, the researcher will distribute questionnaire for the subject of the research in order to gathering the data. The researcher distributed the questionnaire for English Teachers of Senior High School in Kepahiang.

D. Research Instrument

Instrument is a tool or facility that is used to the research to collect data.⁷² It means that the research instrument helped the researcher to get the information

⁷¹ Sugiyono, Op.Cit. p. 94

needed that related to the research. The using of the research instrumental lead the researcher in gathering the information or data needed that related to this research. In this research, the researcher used instruments for helping to complete the data. The research instrument are:

1. Questionnaire Item

According to Nasution, questionnaire is a list of questionnaire as a technique in collecting data that were answered by the respondents themselves.⁷³ All questionnaire is 45 items. Consists of 40 items for test in teaching, it was about how the teachers understand the procedures in designing a test. Then 5 items was about teachers problems in designing a test. In constructing questionnaires, the questions designed based on the theories of design test in teaching and problems in designing test that provided in review of related theories.

Moreover before deciding to distributed this questionnaires to the respondents to ensure the validity of each items in this questionnaires, the researcher used content validity. This kind of validity was that the instrument must show that it fairly and comprehensively covers the domainor items that it purpost to cover.⁷⁴ It means someone who competence in such field could be asking his consideration to evaluate the accuracy of items in questionnaire.

For the first question, the researcher made the item of the questionnaire that refers to teachers problems in designing test. The researcher made in

⁷² Suharsimi Arikunto, *Prosedur Penelitian*, (Jakarta: Rineka Cipta, 2002), p. 136

⁷³ Irwan Nasution, *Metode Penelitian Sosial*, (Jakarta: Rineka Cipta, 1995), p. 65

⁷⁴ Louis Cohen, Lawrence Manion, and Keith Morrison, *Research Method in Education*, (London: Routledge, 2007) p. 137

based on the theory of test in teaching and indicators per each theory. Item of questions can be seen in the table below :

Table 3

Instrument of Questionnaire 1

How do the teachers understand the procedure of all parts in designing a summative test

No	Test in Teaching	Indicators	Item of Statement	Respond			
				A	O	S	N
1	Characteristics of a good test	Test is not excessively expensive,	1. The test I made is easy to conduct and does not require expensive costs.				
		Test stays within appropriate time constraints.	2. In making the test, I consider the completion time of the test.				
		Test is relatively easy to administer and has a scoring or evaluation procedure that is specific and time-efficient.	3. The test I made is easy to check and comes with clear instructions.				
		Test is consistent and dependable	4. When tested on students, the test results can be trusted.				

		Test should yield similar result	5. When the same test is tested on the same student, the test results will have consistent scores.				
		Test valids to the assesment purpose.	6. The test I designed can measure students' abilities accurately.				
		Test are based on a reliable analysis of the skills we want to measure.	7. The test I made is in accordance with the language skills that I want to measure.				
		Authenticity of the features of a target language	8. The test consists of questions that reflect the real conditions of English use.				
		The impact of examinations to students	9. The test results provide feedback on students' abilities.				
		The impact of examinations to students	10. The test results provide feedback on the development of teacher learning.				
2	Planning a test	Identify the assessment goals	11. I identify the assessment's purpose before designing the test.				
		Focus on Standard Competence (SK) and Basic Competence	12. I identify the Competency Standards (SK) and Basic Competencies (KD) before				

		(KD).	designing the test.				
		Identify the test instrument.	13. I identify the type of test I will use before designing the test.				
		Design the examination grid with its suspension guidelines	14. I create a blueprint or scope of questions before designing the test.				
3	Designing a test	The items are matched to the indicators	15. The question items are adjusted to the question indicators.				
		Make sure there is only one answer key for each item.	16. In multiple-choice questions, each question has only one correct answer.				
		The material is matched to the assessment goals.	17. The materials contained in the questions are adjusted to the assessment's purpose.				
		The test materials should consider the urgency.	18. The materials contained in the questions are adjusted to students' abilities.				
		The test materials should consider the relevance.	19. The materials contained in the questions are adjusted to other lesson materials.				
		The test materials	20. The materials contained in the				

		should consider the continuity.	questions are a continuation of previous materials.				
		The test materials should have beneficial in learners' daily	21. The materials contained in the questions consider the concept of usability in everyday life.				
		The material is appropriate to the students' level.	22. The materials contained in the questions are adjusted to the students' levels of education.				
		The distractors are available in the answers.	23. The distractors on the answer choices work well.				
		State the stem of the item in simple, clear language.	24. The subject matters are formulated in a clear language.				
		Present a single clearly formulated problem in the stem of the item.	25. The problems in the question items are clearly formulated.				
		Avoid verbal clues that might enable students to select the correct answer or to eliminate an incorrect alternative.	26. In making question items, I avoid using instruction leading to the correct answer.				

		State the stem of the item in positive form, wherever possible.	27. In making question items, I use questions in the positive form.				
		Emphasize negative wording wherever it is used in the stem of an item.	28. In making question items, I underline or italicize questions with negative statements.				
		Make all alternatives grammatically consistent with the stem of the item.	29. Each answer choice has almost the same formula.				
		Avoid using the alternative "all of the above," and use "none of the above" with the extreme caution.	30. In making question items, I avoid answer choices such as "all answers are correct", "none of the answers are correct", or similar statements.				
		Vary the relative length of the correct answer to eliminate length	31. The length of the answer alternatives is relatively the same, there is no very long or very short answer alternative.				
		Sorted the answer based	32. In making question items, I				

		on the numbers and times.	sort the alternative answers in the form of numbers and time.				
		Use the effective item format.	33. In making question items, I really pay attention to the format of the questions.				
		Make certain each item is independent of the other items in the test.	34. In making question items, I make sure there are no dependencies between one question and another.				
		Present a communicative language.	35. In making question items, I use communicative sentences.				
		The statement present a common language.	36. In making question items, I use a standard language in accordance with the English rules.				
		Present a clear statement to avoid misunderstanding to the item.	37. In making question items, I use sentences without any multiple interpretations.				
		The statement is not containing the offensive	38. In making question items, I avoid using offensive or inappropriate				

		words.	words.				
		The items use a question/comm and that demands the unravelled answers.	39. In making essay questions, I use question words/commands requiring an answer in the form of description.				
		Each items should have its suspension guidelines.	40. In making essay questions, I also prepare a scoring guideline.				

The item of questionnaire above was for teachers procedures in designing a test. As mentions above, the researcher use second questionnaire to answer the second research questions about the factors that cause the problems for the teachersin designing summative test. The process of making the item of questions was same to the steps for the first questionnaire before. The item of questions van be seen in the tabel below :

Table 4

Instrument of Questionnaire 2

Teachers' Problems in Designing Summative Test

No	Problems in designing a test	Indicators	Item of Statement	Respond			
				A	O	S	N
1	The relationship	Class and students	1. I develop a good communication				

	and communication between teachers and other elements outside the teacher	communicative frequently.	with students to avoid the distance between teacher and students.				
		Class management democratization	2. I provide students the freedom to be creative and come up with fresh ideas.				
		Teachers should share and evaluate the assessment strategy with their colleagues.	3. I evaluate my assessment strategy with my colleagues at school.				
		Feedback from students to teachers to understand the students' level of understanding;	4. I use the test results to determine the level of students' understanding of the lesson materials.				
		The support from the institute that facilitates of assessment practices.	5. The school provides facilities to support the making of test questions.				

E. The Technique of Data Analysis

According to Kasirman, analysis means categorizing, ordering, manipulating and summarizing of data to obtain answer to research questions.⁷⁵ In this research, the data has taken by using questionnaire and then it was classified

⁷⁵ Kasirman, *Metodologi Penelitian*, (UIN Malang: Malang Press, 2008), p. 128

into some category, after that the data was analyzed to describe the real condition well. The data was analyzed based on each instrument:

1. The Data from Questionnaire

After researcher get the answer of how teachers understand the procedure of all parts in designing a summative test and about teachers' problems in designing summative test, the data was collected as the teachers response in tabulate, and find out the teachers by using percentage technique. The steps were :

- a. The first steps for analysis the data, the researcher divided the teachers answer in group with same categories. It means for those two questionnaires were classified into each item question to check the teachers respond.
- b. The seconds steps were the researcher accounted the number of teachers given answer in the item based on the questionnaire. In this research, to account the point the researcher used rating scale that has been explained before.
- c. The third steps the researcher used a 4-value Likert Scale to assess the data from the questionnaire, which is used to measure the mean score of teachers' understanding and teachers' problems in designing english summative test. Each statement was created with the value of the mean score in mind (Always = 4, Often = 3, Seldom = 2, Never = 1), the formula is described as follow :

$$x = \frac{\{(f A \times 4) + (f O \times 3) + (f S \times 2) + (f N \times 1)\}}{N}$$

Where :

F : Frequency

A : Always

O : Often

S : Seldom

N : Never

- d. The fourth steps were the researcher accounted the percentage of the each items of questionnaire based on the teachers answer in questionnaire. It follows the formula :⁷⁶

$$P = \frac{\sum score}{Max score} \times 100\%$$

Where :

P : Percentage

\sum score : Total score

Max score : Maximum score

The teachers' understanding were sorted into four categories using Linkert scales. The outcome of the calculation was sorted into good and poor understanding in designing english summative test based on the accounts with four categories. The results of the questionnaire revealed the presentage, which was then classified into two categories:

⁷⁶ Arif, *Cara Menghitung Skala Likert* at <http://naufansapoetra.blogspot.co.id/2015/11/cara-menghitung-kuesioner-skala-likert.html>. Retrieved on Monday, 6 September 2021

1) > 51,25% =Good

2) < 51,25% =Poor

- e. The fifth steps the researcher accounted the problems faced by teachers based on each items of questionnaire and seen the higher percentage of each procedure.

As follow the teachers' understanding, the teachers' problems were also sorted into four categories using Linkert scales. The outcome of the calculation was sorted into good and poor understanding based on the accounts with four categories. The results of the questionnaire revealed the presentage, which was then classified into two categories:

- 1) > 51,25% =Good (teachers' didnt have any problems in designing test)
- 2) < 51,25% =Poor (Teachers had problems in designing test)

CHAPTER IV

FINDING AND DISCUSSION

This chapter provides the findings and discussion of this research, in which the given details are presented in accordance with the focuses explained in the problem statements of this research.

A. Finding

The finding of this research was presented on the basis of two research questions formulated as the primary orientation of this research. As absorbed from the core of research questions, the data orientation comprises 1) teachers' understanding of the procedures of designing a summative test, and 2) teachers' problems in designing a summative test

1. Teachers' Understanding of the Procedures of Designing a Summative Test

A questionnaire distributed online by using *Google Form* to all the research samples, 62 English teachers, was employed to obtain all the necessary data to determine the teachers' understanding of the procedures of designing a summative test. According to the data garnered from the questionnaire, there were varied points to be portrayed with respect to the levels of teachers' understanding of a summative test design. The following conceptual table was presented to display the overall condition of the data.

Table 5

Teachers' Understanding of the Procedures of Designing a Summative Test

Indicators	Statements		Total Number of Teachers' Response								Average	Total Average
			A		O		S		N			
	No	Items of Statement	F	%	F	%	F	%	F	%	%	%
Characteristics of a good test	1	Test is not excessively expensive	34	54,8	18	29	10	16,2	-		84,6	80,85
	2	Test stays within appropriate time constraints.	22	35,5	28	45,2	10	16,1	2	3,2	78,2	
	3	Test is relatively easy to administer and has a scoring or evaluation procedure that is specific and time-efficient.	26	41,9	26	41,9	9	14,5	1	1,6	81	
	4	Test is consistent and dependable	37	59,7	16	25,8	9	14,5	-		85,1	
	5	Test should yield similar result	23	37,1	29	46,8	8	12,9	2	3,2	83,5	
	6	Test valids to the assesment purpose.	25	40,3	23	37,1	14	22,6	-		81	
	7	Test are based on a reliable analysis of the skills we want to measure.	24	38,7	25	40,3	13	21	-		80,4	
	8	Authenticity of the features of a target language	25	40,3	26	41,9	9	14,5	2	3,2	79,8	
	9	The impact of examinations to students	28	45,2	25	40,3	8	12,9	1	1,6	82,3	
	10	The impact of examinations to teachers	19	30,6	21	33,9	19	30,6	3	4,8	72,6	
Planning a test	11	Identify the assessment goals	25	40,3	22	35,5	15	24,2	-		79	79,4

	12	Focus on Standard Competence (SK) and Basic Competence (KD).	22	35,5	26	41,9	14	22,6	-		78,2	
	13	Identify the test instrument.	28	45,2	20	32,3	13	21	1	1,6	80,2	
	14	Design the examination grid with its suspension guidelines	25	40,3	25	40,3	12	19,4	-		80,2	
Designing a test	15	The items are matched to the indicators	28	45,2	25	40,3	9	14,5	-		82,7	80,53
	16	Make sure there is only one answer key for each item.	30	48,4	18	29	14	22,6	-		81,5	
	17	The material is matched to the assessment goals.	24	38,7	25	40,3	12	19,4	1	1,6	79	
	18	The test materials should consider the urgency.	24	38,7	25	40,3	13	21	-		80,4	
	19	The test materials should consider the relevance.	19	30,6	29	46,8	11	17,7	3	4,8	75,8	
	20	The test materials should consider the continuity.	31	50	21	33,9	10	16,1	-		83,5	
	21	The test materials should have beneficial in learners' daily	28	45,2	24	38,7	9	14,5	1	1,6	81,9	
	22	The material is appropriate to the students' level.	24	38,7	31	50	6	9,7	1	1,6	81,5	
	23	The distractors are available in the answers.	23	37,1	23	37,1	15	24,1	1	1,6	77,4	
	24	State the stem of the item in simple, clear language.	23	37,1	25	40,3	13	21	1	1,6	84,7	

25	Present a single clearly formulated problem in the stem of the item.	23	37,1	22	35,5	16	25,8	1	1,6	77
26	Avoid verbal clues that might enable students to select the correct answer or to eliminate an incorrect alternative.	31	50	22	35,5	9	14,5	-		83,9
27	State the stem of the item in positive form, wherever possible.	23	37,1	25	40,3	14	22,6	-		78,6
28	Emphasize negative wording wherever it is used in the stem of an item.	29	46,8	23	37,1	10	16,1	-		82,7
29	Make all alternatives grammatically consistent with the stem of the item.	28	45,2	23	37,1	11	17,7	-		81,9
30	Avoid using the alternative "all of the above," and use "none of the above" with the extreme caution.	17	27,4	26	41,9	17	27,4	2	3,2	82
31	Vary the relative length of the correct answer to eliminate length	32	52,6	20	32,3	10	16,1	-		73,4
32	Sorted the answer based on the numbers and times.	22	35,5	29	46,8	11	17,7	-		79,4
33	Use the effective item format.	25	40,3	29	46,8	7	11,3	1	1,6	81,5
34	Make certain each item is independent of the other items in the test.	28	45,2	28	45,2	6	9,7	-		83,9

35	Present a communicative language.	22	35,5	25	40,3	14	22,6	1	1,6	77,4
36	The statement present a common language.	23	37,1	22	35,5	17	27,4	-		77,4
37	Present a clear statement to avoid misunderstanding to the item.	29	46,8	24	38,7	9	14,5	-		83,1
38	The statement is not containing the offensive words.	28	45,2	27	43,5	5	8,1	2	3,2	80,6
39	The items use a question/command that demands the unravelled answers.	32	51,6	21	33,9	9	14,5	-		84,3
40	Each items should have its suspension guidelines.	26	41,9	18	29	18	29	-		78,2

Notes:

A = Always

O = Often

S = Seldom

N = Never

The conceptual table above indicated that most English teachers involved as the samples of this study already had a good level of understanding of the procedures of designing a summative test. The levels of understanding in question were portrayed on their responses toward each questionnaire item presented on the above table, in which the items were classified based on several themes or indicators taken from the constructed theories, covering 1) characteristics of a good test, 2) planning a test, and 3) designing a test.

In regards to the characteristics of a good test, based on the responses to the related questionnaire items, most teachers showed that they already had a good understanding of the items comprising a good test's characteristics, namely 1) test is not excessively expensive, 2) test stays within appropriate time constraints, 3) test is relatively easy to administer and has a scoring or evaluation procedure that is specific and time-efficient, 4) test is consistent and dependable, 5) test should yield similar result, 6) test is valid to the assesment purpose, 7) test is based on a reliable analysis of the skills to be measured, 8) authenticity of the features of a target language, 9) the impact of examinations to students, and 10) the impact of examinations to teachers. Correspondingly, based on the responses gathered using the questionnaire, most teachers stated that they made the test easy to conduct and did not require expensive costs (54.8% always; 29% often), considered the completion time of the test (35.5% always; 45.2% often), made the test easy to check and came with clear instructions (41.9% always; 41.9% often), made the test with trusted results (59.7% always; 25.8% often), made the test with consistent scores (37.1% always; 46.8% often), designed the test capable to measure students' abilities accurately (40.3% always; 37.1% often), made the test per the language skills to be measured (38.7% always; 40.3% often), made the test reflecting the real conditions of English use (40.3% always; 41.9% often), made the test which could provide feedback on students' abilities (45.2% always; 40.3% often), and made the test which could provide feedback on teachers' development (30.6% always; 33.9% often).

Afterwards, with respect to the indicator of planning a test, most teachers' responses indicated that they already understood well about the items comprising a test planning, namely 1) identify the assessment goals, 2) focus on Standard Competence (*SK*) and Basic Competence (*KD*), 3) identify the test instrument, and 4) design the examination grid with its suspension guidelines. Accordingly, based on the responses to the questionnaire items, most teachers stated that they identified the assessment's purpose before designing the test (40.3% always; 35.5% often), identified the Competency Standards (*SK*) and Basic Competencies (*KD*) before designing the test (35.5% always; 41.9% often), identified the type of test to be used before designing the test (45.2% always; 32.3% often), and created a blueprint or scope of questions before designing the test (40.3% always; 40.3% often).

Furthermore, regarding the indicator of designing a test, most teachers were found to have a good understanding in relation to the corresponding items, namely 1) the items are matched to the indicators, 2) make sure there is only one answer key for each item, 3) the material is matched to the assessment goals, 4) the material is matched to the assessment goals, 5) the test materials should consider the urgency, 6) the test materials should consider the relevance, 7) the test materials should consider the continuity, 8) the test materials should have beneficial in learners' daily, 9) the material is appropriate to the students' level, 10) the distractors are available in the answers, 11) state the stem of the item in simple, clear language, 12) present a single clearly formulated problem in the stem of the item, 13) avoid verbal

clues that might enable students to select the correct answer or to eliminate an incorrect alternative, 14) state the stem of the item in positive form, wherever possible, 15) emphasize negative wording wherever it is used in the stem of an item, 16) make all alternatives grammatically consistent with the stem of the item, 17) avoid using the alternatives “all of the above” and “none of the above” with the extreme caution, 18) vary the relative length of the correct answer to eliminate length, 19) sort the answer based on the numbers and times, 20) use the effective item format, 21) make certain each item is independent of the other items in the test, 22) present a communicative language, 23) the statement present a common language, 24) present a clear statement to avoid misunderstanding to the item, 25) the statement is not containing the offensive words, 26) the items use a question/command that demands the unravelled answers, and 27) each item should have its suspension guidelines.

As regards to the above indicator, designing a test, the responses on the questionnaire items illustrated that most teachers designed the test with the following criteria: the question items were adjusted to the question indicators (45.2% always; 40.3% often), in multiple-choice questions, each question had only one correct answer (48.4% always; 29% often), the materials contained in the questions were adjusted to the assessment's purpose (38.7% always; 40.3% often), the materials contained in the questions were adjusted to students' abilities (38.7% always; 40.3% often), the materials contained in the questions were adjusted to other lesson materials (30.6% always; 46.8%

often), the materials contained in the questions were a continuation of previous materials (50% always; 33.9% often), the materials contained in the questions considered the concept of usability in everyday life (45.2% always; 38.7% often), the materials contained in the questions were adjusted to the students' levels of education (38.7% always; 50% often), the distractors on the answer choices worked well (37.1% always; 37.1% often), the subject matters were formulated in a clear language (37.1% always; 40.3% often), the problems in the question items were clearly formulated (37.1% always; 35.5% often), avoided using instruction leading to the correct answer (50% always; 35.5% often), used questions in the positive form (37.1% always; 40.3% often), underlined or italicized questions with negative statements (46.8% always; 37.1% often), each answer choice had almost the same formula (45.2% always; 37.1% often), avoided answer choices such as "all answers are correct", "none of the answers are correct", or similar statements (27.4% always; 41.9% often), the length of the answer alternatives was relatively the same (51.6% always; 32.3% often), sorted the alternative answers in the form of numbers and time (35.5% always; 46.8% often), paid attention to the format of the questions (40.3% always; 46.8% often), made sure there were no dependencies between one question and another (45.2% always; 45.2% often), used communicative sentences (35.5% always; 40.3% often), used a standard language in accordance with the English rules (37.1% always; 35.5% often), used sentences without any multiple interpretations (46.8% always; 38.7% often), avoided using offensive or inappropriate words (45.2% always; 43.5%

often), used question words/commands requiring an answer in the form of description in essay questions (51.6% always; 33.9% often), and prepared a scoring guideline for essay questions (41.9% always; 29% often).

2. Teachers' Problems in Designing a Summative Test

Based on the data obtained using the online questionnaire, there were several problems experienced by the English teachers in designing a summative test. The following conceptual table was presented to display the overall condition of the data.

Table 6
Teachers' Problems in Designing a Summative Test

Indicator		Items of Statement	Total Number of Teachers' Response								Average
			A		O		S		N		
			F	%	F	%	F	%	F	%	
The relationship and communication between teachers and other elements outside the teacher	1	Class and students communicative frequency	4	6,45	7	11,3	28	45,2	23	37,1	46,8
	2	Class management democratization	15	24,2	16	25,8	26	41,9	5	8,2	66,5
	3	Teachers should share and evaluate the assessment strategy with their colleagues	17	27,4	15	24,2	30	48,4	-		69,8
	4	Feedback from students to teachers to understand the students' level of understanding	4	6,45	5	8,1	26	41,9	27	43,5	44,4
	5	The support from the institute that facilitates the assessment practices	17	27,4	18	29	27	43,5	-		71

Notes:

A = Always

O = Often

S = Seldom

N = Never

The conceptual table above indicated that there were several obstacles faced by most teachers in designing a summative test. However, there was a slight difference in the percentages of the total responses. Teachers' problems in question were portrayed on the responses to the questionnaire items classified based on the theme or indicator garnered from the constructed theories, namely the relationship and communication between teachers and other elements outside the teacher. In regards to the aforementioned indicator, most teachers showed that they experienced some problems in several aspects, namely 1) class and student communication frequency and 2) feedback from students to teachers to understand the students' level of understanding. Meanwhile, nearly half of the total number of teachers faced the obstacles in several other aspects, namely 1) class management democratization, 2) assessment strategy sharing and evaluation with colleagues, and 3) the support from the institute that facilitates the assessment practices.

Accordingly, based on the responses to questionnaire items, 53.2% of teachers stated that they did not develop a good communication with students to avoid the distance between teacher and students, 61.3% admitted that they did not use the test results to determine the level of students' understanding of the lesson materials, 48.4% revealed that they did not provide students the freedom to be creative and come up with fresh ideas, 48.4% stated that they

did not evaluate the assessment strategy with their colleagues at school, and 29% admitted that the school did not provide facilities to support the making of test questions.

B. Discussion

This section discusses all the data obtained, displayed, and analyzed in prior. There were two major scopes of findings being addressed, namely 1) teachers' understanding of the procedures of designing a summative test and 2) teachers' problems in designing a summative test.

As obtained from the questionnaire, it could be seen that most English teachers involved as the research samples had a good level of understanding of the procedures of designing a summative test, including the characteristics of a good test, planning a test, and designing a test. Thus, the teachers' levels of understanding were considered to be in line with the constructed theories about the procedures of designing a summative test. Accordingly, based on Brown's theories regarding the characteristics of a good test, by having a good level of understanding, the English teachers were believed to be able to make the tests without neglecting the aspects of practicality⁷⁷, reliability⁷⁸, validity⁷⁹, authenticity⁸⁰, and washback⁸¹. In addition, by having a well comprehension in planning a test, the English teachers were considered capable of creating the question items by employing the instructions provided

⁷⁷ H. Douglas Brown, *Language Assessment Principles and Classroom Practices*, (Longman : London, 2004), p. 19

⁷⁸ Ibid, p. 20

⁷⁹ Ibid, p. 22

⁸⁰ Ibid, p. 28

⁸¹ Ibid

by Depdiknas⁸², namely identifying the assessment goals, focusing on Standard Competence (*SK*) and Basic Competence (*KD*), identifying the test instrument, and designing the examination grid with its suspension guidelines. Furthermore, the teachers' responses indicated that they also had understood well about designing a test, which in turn made them able to provide the test items in accordance with the previously mentioned rules in designing test items proposed by *Departemen Pendidikan Nasional*⁸³.

Nevertheless, despite the good level of understanding possessed by the teachers, there were still some problems they had to struggle with in designing a summative test. Most English teachers involved in the present study faced some obstacles regarding the relationship and communication between teachers and other elements outside the teacher. Thus, it somehow implied that instead of coming from within, the problems emerged from the surrounding environment influencing the process of designing a summative test. There were five possible problems that the teachers might experience in designing a summative test, including class and student communication frequency, class management democratization, assessment strategy sharing and evaluation with colleagues, feedback from students to teachers to understand the students' level of understanding, and the support from the institute that facilitates the assessment practices⁸⁴. According to teachers' responses to questionnaire

⁸² Departemen Pendidikan Nasional, *Panduan Penulisan Butir Soal*, (Jakarta: Depdiknas, 2008), p. 7-8

⁸³ Departemen Pendidikan Nasional, *Pedoman Umum pengembangan Penilaian*, (Jakarta: Depdiknas, 2003), p. 76-80

⁸⁴ As Homby, *Problems In The Delivery Of The School-Based Psychoeducational Services To The Asian Immigrant Children*, pp. 89-90

items, it could be seen that most teachers faced two out of five aforementioned problems, namely 1) class and student communication frequency, in which the teachers admitted that they could not develop a good communication with students to avoid the distance between teacher and students, and 2) feedback from students to teachers to understand the students' level of understanding, in which the teachers did not use the test results to determine the level of students' understanding of the lesson materials.

Furthermore, without disregarding the slight difference of the total percentage of teachers' responses, almost half of the total number of teachers to some extent had to struggle with the other three problems in designing a summative test, namely 1) class management democratization, 2) assessment strategy sharing and evaluation with colleagues, and 3) the support from the institute that facilitates the assessment practices, which somehow affected the teachers' performances in designing a proper test or question items in spite of their good understanding of the procedures of designing a summative test.

CHAPTER V

CONCLUSION AND SUGGESTION

In this chapter, the researcher concluded the research results by answering the research questions and giving some points of suggestion for the parties involved.

A. Conclusion

Based on finding and discussion in the previous chapter, the researcher finally concluded that:

1. 80,3 % teachers involved as the samples of this study already had a good level of understanding of the procedures of designing a summative test. The teachers' levels of understanding were portrayed on their responses to the questionnaire items classified based on several themes or indicators taken from the constructed theories, including 1) 80,85% teachers understand of characteristics of a good test, 2) 79,4 % teachers understand of planning a test, and 3) and 80,53 % teachers understand of designing a test.
2. There were several obstacles faced by most teachers in designing a summative test. The teachers' problems were classified based on the theme or indicator garnered from the constructed theories, namely the relationship and communication between teachers and other elements outside the teacher. In regards to the aforementioned indicator, most teachers showed that they experienced some problems in several aspects, 1) 53,2% teachers had

problems with class and student communication frequency and 2) 61,3 % teachers did not use feedback from students to teachers to understand the students' level of understanding. Meanwhile, despite a slight difference in the total percentage of responses, nearly half of the total number of teachers faced the obstacles in several other aspects, namely 1) 51,6% teachers provided a the freedom class management democratization, 2) 51,8% teachers shared and evaluated assessment strategy with colleagues, and 3) 71% teachers admitted the school the support that facilitates assessment practices.

B. Suggestion

Based on the present study results and conclusion, some suggestions were given to several parties involved in English teaching and learning, namely English teachers, school stakeholders, and other researchers.

1. English teachers

Based on the study findings, most English teachers were found to possess a good level of understanding of the procedures of designing a summative test. However, in spite of their comprehension, most teachers could not be completely free from all the possible problems in designing a summative test. The research findings revealed that the teachers had to struggle with the issues triggered by the relationship and communication between teachers and other elements outside the teacher, particularly on class and student communication frequency as well as feedback from students to teachers to understand the students' level of understanding. Therefore, the researcher suggested the English teachers to try developing

a good communication with the students to decrease the gaps or distance emerged between teacher and student in the classroom. In addition, the teachers were suggested to take into account the summative test results in determining their students' levels of understanding, so that the students' learning outcomes could be described properly.

2. School stakeholders in Indonesia

The results of this research at some point ended up with a suggestion for the school stakeholders in Indonesia to start carrying out an improvisation regarding the support for teachers in form of the facilities they need in designing a summative test. Thus, the students' learning outcomes could be determined without any significant obstacles involving the lack of support of media or other related supporting facilities. Despite the fact that most teachers did not experience the issues regarding the facilities provided by the schools, the school stakeholders should not neglect the slight difference in the total percentage of teachers' responses to the related questionnaire items, so that the process of summative test design would still be carried out properly in the future.

3. Other researchers

With respect to the depth of the present study, this research was only delimited on probing into English teachers' understanding of the procedures of designing a summative test and also their problems in designing a summative test by administering online questionnaires,

leading to the research data which were limited to the closed responses given by the teachers. Therefore, the future researchers were suggested to employ the other research instruments such as interview or observation in order to explore more data, so that the influencing factors and teachers' in-depth explanation regarding the research topic could be discovered.

REFERENCE

Alberta. *Formative and Summative Assessment*. (Chicago: Northern Illinois University of Chicago Faculty Development and Instructional Design Centre, 2008). Unpublished. Retrived from Wwww. HOI science.com, on August 20th 2018

Arif, *Cara Menghitung Skala Likert at http://naufansapoetra.blogspot.co.id/2015/11_cara-menghitung-kuesioner-skala-likert.html*. Retrieved on Monday, 6 September 2021

Arikunto Suharsimi, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: Renika, 2009), p. 128

Athur Hughes, *Testing for Language Teacher*, (Great Britain: Cambridge University Press., 2003), p.13

As Homby, *Problems In The Delivery Of The School-Based Psychoeducational Services To The Asian Immigrant Children*, pp. 89-90

B. S. Bloom, *Taxonomy of Educational Objectives*, (New York: Longman, 1956), p. 48-50

Bostwick, R. M and Gakuen K, *Evaluating young EFL Learners : Problems and Solutions*, In Brown, J. D. And Yamashita,S. O. (eds), *JALT Allied Materials Language Testing in Japan*, Tokyo: The japan Association for Language Teaching. p.57

David P. Harris, *Testing English as a Second Language*, (New York: McGraw Hill Book Company, 1969), p. 14

Departemen Pendidikan Nasional, *Panduan Penulisan Butir Soal*, (Jakarta: Depdiknas, 2008), p. 7-8

Departemen Pendidikan Nasional, *Pedoman Umum pengembangan Penilaian*, (Jakarta: Depdiknas, 2003), p. 76-80

Desri Susiyanti, 2019, *Theachers' Difficulties in Theaching English as a Foreign Language at SMK Muhammadiyah 3 Makassar*

Donal Ary, *Introduction to Research in Education*, (USA: Wadsworth Group A Division Of Thompson Learning Inc, 2010), 8th Ed., p.148

Everlyn Olouch, *Challenges Faced by Tutors in Setting of Examinations*, Journal of education, Vol. 5, No. 17, 2014

Fu Ren, *The Construction of a Harmonious Relationship Between Teachers and Students to Improve the Actual Effect of Moral Education*, Journal of Social Science, Vol. 11, No. 5, p. 13-17

Glenn Fulcher, *Practical Language Testing*, (Britain: Hodder Education, An Hachette UK Company, 2010), p.4.

Grant Henning, *A Guide to Language Testing: Development, Evaluation and Research*, (China: Heinle & Heinle Publisher, 2001), p.6

H. Douglas Brown, *Language Assessment Principles and Classroom Practices*, (Longman : London, 2004), p. 19

H. Douglas Brown, *Teaching by Principles; An Interactive Approach to Language Pedagogy*, (New York: A Pearson Education Company, 2015), 4th. Ed., p. 500

Irwan Nasution, *Metode Penelitian Sosial*, (Jakarta: Rineka Cipta, 1995), p. 65

Isabell Sengkaton, 2020, *An Analysis Of Teacher-Made Summative Test of English of SMA Negeri 3 Palembang*

James Dean Brown, *Testing in Language Programs*, (New Jersey: Prentice Hall Regents, 1996), p.10

Joshua A. Russel & James R. Austin, *Assessment Practices of Secondary Music Teachers*, *Journal of Research in Music Education*, Vol. 58 No. 1, p. 37-38

Kemal Isci, *Internal and External Factors affecting Teachers Adoption of Formative Assessment to Support learning*, *International Journal of Social, Educational, Economic, Business, and Industrial Engineering*, Vol. 10 No. 8, 2016, p. 4-8

Louis Cohen, Lawrence Manion, and Keith Morrison, *Research Method in Education*, (London: Routledge, 2007) p. 137

M. Aries Taufiq, et.all., 2029, *Item Analysis and Teachers' Factors in Designing a Test*

Nasution, *Metode Research*, (Jakarta: Bumi Aksara, 1998), p. 101

N. E. Gronlund, *Measurement and Evaluation in Teaching, Fourth Edition*, (New York: MacMillan, 1981), p. 178

Norman E. Gronlund, *Constructing Achievement Tests. Second Edition*, (USA: Prentice-Hall Inc., 1977), p. 57

Nurul Zuriyah, *Metodology Sosial dan pendidikan, Teori Aplikasi*, (Jakarta: Bumi Aksara, 2006), p. 37

Nurhalimah & Fahriany, *Determining the Quality of English Teacher-Made Test: How Excellent is Excellent?*, Journal of Linguistics, and Literature, Vol. 5, No. 1, 2019

Paivi Atjonen, *Teacher's View of Their Assessment Practice*, The Curriculum Journal, Vol. 25 No. 2, 2014, p. 243

Purwanto, *Metodologi Penelitian Kuantitatif Untuk Psikologi Dan Pendidikan*, (Yogyakarta: Pustaka Pelajar, 2008), p. 226

Richard I. Arends, *Learning to Teach*, (New York: McGraw Hills, 2012), 10th. Ed., p.221

RL. Thorndike and Hagen E., *Measurement and Evaluation in Psychology and Education*, 2nd Ed, (USA: G and C Meriam Co, 1962), p. 47

Syukarman Syarnubi, *Metodology penelitian Kuantitatif dan Kualitatif*, (Curup: LP2 STAIN Curup, 2011), p. 102

Sugiono, *Metode Penelitian Kuantitatif Kualitatif dan R&D*, (Bandung: Alfabeta CV, 2010), p. 94

T. Fatimah Djajasudrama, *Metode Linguistic, Ancangan Metode Penelitian dan kajian*, (Bandung: PT Eresco, 1993), p. 3

Wilmar Tinambunan, *Evaluasi Hasil Belajar*, (Jakarta: Departemen Pendidikan dan Kebudayaan, 1988) p.34

Walter Dick, Lou Carey, James O. Carey, *The Systematic Design of Instruction*, (USA: Library of Congress Cataloging, 2015) , p.6-8

Wynne Harlen, *Teachers' Summative Practices and Assessment for Learning-Tension and Synergies*, *The Curriculum Journal*, Vol. 16 No. 2, 2005, p.214

Appendix

No	Indicator	Sub-Indicator	Total Number of Teachers' Response			
			A	O	S	N
1	Characteristics of a good test	Test is not excessively expensive	34	18	10	-
		Test stays within appropriate time constraints.	22	28	10	2
		Test is relatively easy to administer and has a scoring or evaluation procedure that is specific and time-efficient.	26	26	9	1
		Test is consistent and dependable	37	16	9	-
		Test should yield similar result	23	29	8	2
		Test valids to the assesment purpose.	25	23	14	-
		Test are based on a reliable analysis of the skills we want to measure.	24	25	13	-
		Authenticity of the features of a target language	25	26	9	2
		The impact of examinations to students	28	25	8	1
		The impact of examinations to teachers	19	21	19	3
		2	Planning a test	Identify the assessment goals	25	22
Focus on Standard Competence (SK) and Basic Competence (KD).	22			26	14	-
Identify the test instrument.	28			20	13	1
Design the examination grid with its suspension guidelines	25			25	12	-

3	Designing a test	The items are matched to the indicators	28	25	9	-
		Make sure there is only one answer key for each item.	30	18	14	-
		The material is matched to the assessment goals.	24	25	12	1
		The test materials should consider the urgency.	24	25	13	-
		The test materials should consider the relevance.	19	29	11	3
		The test materials should consider the continuity.	31	21	10	-
		The test materials should have beneficial in learners' daily	28	24	9	1
		The material is appropriate tho the students' level.	24	31	6	1
		The distractors are available in the answers.	23	23	15	1
		State the stem of the item in simple, clear language.	23	25	13	1
		Present a single clearly formulated problem in the stem of the item.	23	22	16	1
		Avoid verbal clues that might enable students to select the correct answer or to eliminate an incorect alternative.	31	22	9	-
		State the stem of the item in positive form, wherever possible.	23	25	14	-
		Emphasize negative wording wherever it is used in the stem of an item.	29	23	10	-
		Make all alternatives gramatically consistent with the stem of the item.	28	23	11	-
		Avoid using the alternative "all of the above," and use	17	26	17	2

		“none of the above” with the extreme caution.				
		Vary the relative length of the correct answer to eliminate length	32	20	10	-
		Sorted the answer based on the numbers and times.	22	29	11	-
		Use the effective item format.	25	29	7	1
		Make certain each item is independent of the other items in the test.	28	28	6	-
		Present a communicative language.	22	25	14	1
		The statement present a common language.	23	22	17	-
		Present a clear statement to avoid misunderstanding to the item.	29	24	9	-
		The statement is not containing the offensive words.	28	27	5	2
		The items use a question/command that demands the unravelled answers.	32	21	9	-
		Each items should have its suspension guidelines.	26	18	18	-