

**AN ANALYSIS OF ENGLISH SEMESTER TESTS MADE BY TEACHERS  
FOR GRADE VIII SEKOLAH MENENGAH PERTAMA (SMP) NEGERI  
AT KEC. TAMBANG KAB. KAMPAR RIAU**

**THESIS**



**M. ARIES TAUFIQ  
NIM. 14178065**

*Submitted to fulfill one of the requirements to obtain  
a degree of Master of Education*

**ENGLISH LANGUAGE EDUCATION STUDY PROGRAM  
FACULTY OF LANGUAGES AND ARTS  
UNIVERSITAS NEGERI PADANG  
2016**



## ABSTRAK

**M. Aries Taufiq. 2016. "An Analysis of English Semester Tests Made by Grade VIII Teachers at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau". *Tesis. Pendidikan Bahasa Inggris Program Pascasarjana. Universitas Negeri Padang.***

Tes merupakan sebuah instrumen yang digunakan untuk mengukur hasil belajar siswa setelah mengikuti proses belajar Bahasa Inggris pada rentang waktu tertentu. Penelitian ini bertujuan untuk mengevaluasi soal ujian semester Bahasa Inggris yang dibuat oleh guru kelas VIII. Penelitian ini digolongkan ke dalam penelitian evaluasi. Penelitian ini difokuskan pada validitas, reliabilitas, analisis butir serta faktor dari guru yang mempengaruhi kualitas soal yang dibuat oleh guru kelas VIII Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau. Data dalam penelitian ini yaitu soal ujian semester yang digunakan di SMP Negeri (1, 2, 4 and 6) Kec. Tambang, lembar jawaban siswa dan transkrip wawancara dengan guru di sekolah tersebut. Hasil penelitian ini menunjukkan bahwa soal ujian semester yang dibuat oleh guru kelas VIII perlu direvisi. Dari hasil analisis data ditemukan bahwa soal ujian semester pada masing-masing sekolah masih memiliki validitas yang rendah meskipun hampir semua kompetensi dasar dapat ditemukan dalam soal ujian semester tersebut. Kemudian, koefisien reliabilitas soal juga masih rendah (0.32 untuk SMPN 1, 0.41 untuk SMPN 2, 0.71 untuk SMPN 4, and 0.42 untuk SMPN 6). Di samping itu, berdasarkan analisis butir, pada tingkat kesulitan ditemukan bahwa soal ujian semester yang digunakan didominasi oleh butir sedang. Dalam hal daya pembeda, butir-butir tersebut tidak dapat membedakan siswa berkemampuan tinggi dan rendah. Kemudian, semua soal semester didominasi oleh butir yang berfungsi sebagai pengganggu. Hasil penelitian ini juga menunjukkan bahwa soal ujian semester yang dibuat oleh guru SMPN 4 memiliki kualitas yang lebih baik dibandingkan soal-soal yang digunakan di SMPN 1, 2 dan 6. Hal ini terjadi karena guru di SMPN 4 sering mengikuti pelatihan dalam pembuatan tes. Oleh karena itu, pelatihan dalam pembuatan tes secara signifikan mempengaruhi kualitas soal yang dibuat oleh guru. Penelitian ini menyarankan agar pihak SMP Negeri di Kec. Tambang memberikan kesempatan kepada guru-gurunya untuk mengikuti pelatihan secara berkelanjutan agar kualitas soal yang mereka buat dapat ditingkatkan.

## ABSTRACT

**M. Aries Taufiq. 2016. "An Analysis of English Semester Tests Made by Grade VIII Teachers at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau". *Thesis*. English Graduate Program. Universitas Negeri Padang.**


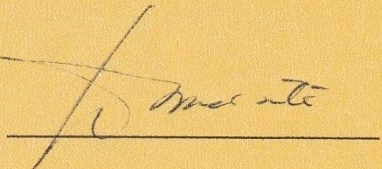
A test is an instrument to measure students' achievement after attending English learning process in certain period of time. This research was aimed to find out and analyze the English semester tests made by grade VIII teachers. This research was classified into evaluation research. It was focused on the validity, reliability, item analysis and factors from teachers that influence the quality of the semester tests made by grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau. The data of this research were the semester tests used at SMP Negeri (1, 2, 4 and 6) Kec. Tambang, students' answer sheets and the transcription from the interview with the teachers. The findings showed that the semester tests made by the teachers need to be revised. From the data analysis, it was found that the tests of each school had low validity although it covered almost all of the basic competences for grade VIII students. Then, the reliability coefficient of the tests was also low (0.32 for SMPN 1, 0.41 for SMPN 2, 0.71 for SMPN 4, and 0.42 for SMPN 6). From the item analysis, based on the level of difficulty it was revealed that the tests were dominated by moderate items. For the discrimination indices, they were not able to discriminate high and low students. In addition, all kinds of semester tests were dominated by items which functioned well. The result of this research showed that the semester test used at SMPN 4 has better quality compared to semester tests used at SMPN 1, 2 and 6 since the teacher at SMPN 4 often join several trainings on test construction. Therefore, training on test construction contributes significantly to the quality of the test. This research suggested the principles of SMP Negeri at Kec. Tambang provide the teachers with continuous trainings especially on test construction.



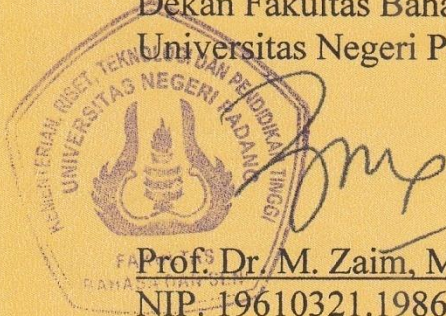
## PERSETUJUAN AKHIR TESIS

---

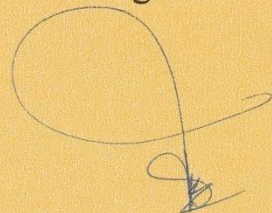
Mahasiswa : **M. Aries Taufiq**  
NIM : 14178065  
Program Studi : Pendidikan Bahasa Inggris

Nama	Tanda Tangan	Tanggal
<u>Prof. Dr. Mukhaiyar, M.Pd.</u> Pembimbing I		12/8-016
<u>Dr. Desmawati Radjab, M.Pd.</u> Pembimbing II		12/8-018

Dekan Fakultas Bahasa dan Seni  
Universitas Negeri Padang

  
Prof. Dr. M. Zaim, M.Hum.  
NIP. 19610321.198602.1.001


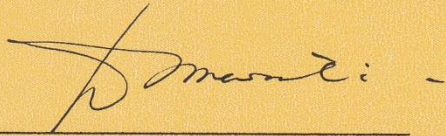
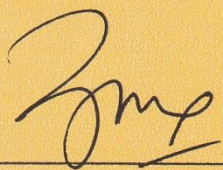

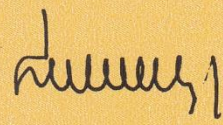
Ketua Program Studi

  
Prof. Dra. Yenni Rozimela, M.Ed., Ph.D.  
NIP. 19620919.198602.1.001



**PERSETUJUAN KOMISI  
UJIAN TESIS MAGISTER KEPENDIDIKAN**

---

No.	Nama	Tanda Tangan
1.	<u>Prof. Dr. Mukhaiyar, M.Pd.</u> (Ketua)	 _____
2.	<u>Dr. Desmawati Radjab, M.Pd.</u> (Sekretaris)	 _____
3.	<u>Prof. Dr. M. Zaim, M.Hum.</u> (Anggota)	 _____
4.	<u>Dr. Refnaldi, M.Litt.</u> (Anggota)	 _____
5.	<u>Prof. Dr. Gusril, M.Pd.</u> (Anggota)	 _____

Mahasiswa

Mahasiswa	: <b>M. Aries Taufiq</b>
NIM	: 14178065
Program Studi	: Pendidikan Bahasa Inggris
Tanggal Ujian	: 10 - 8 - 2016



## SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Karya tulis saya, dengan judul **“An Analysis of English Semester Tests Made by Grade VIII Teachers at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau”** adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik baik di Universitas Negeri Padang maupun di perguruan tinggi lainnya.
2. Karya tulis ini murni gagasan, penilaian, dan rumusan saya sendiri, tanpa bantuan tidak sah dari pihak lain, kecuali arahan tim pembimbing.
3. Di dalam karya tulis ini tidak terdapat hasil karya atau pendapat yang telah ditulis atau dipublikasikan orang lain kecuali dikutip secara tertulis dengan jelas dan dicantumkan sebagai acuan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya, dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan norma dan ketentuan hukum yang berlaku.

Padang, Agustus 2016



M. Aries Taufiq  
NIM. 14178065

## ACKNOWLEDGEMENT

Bismillahirrahmanirrahim. First of all, I would like to express all praise, honor, gratitude, and glory to the Almighty Allah SWT, who has given me guidance and strengths to finish this thesis, entitled **An Analysis of English Semester Tests Made by Grade VIII Teachers at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau**. This thesis was completed as one of the requirements to get Master degree in English Graduate Program, Universitas Negeri Padang.

To finish this thesis, I got a lot of helpful contributions from special people. First, I would like to thank the advisors: Prof. Dr. Mukhaiyar, M.Pd and Dr. Desmawati Radjab, M.Pd, who have given great deal of contributions, valuable advices, meaningful knowledge, time, and assistance in finishing this thesis. Then, special acknowledgements are addressed to Prof. Dr. M. Zaim, M.Hum, Dr. Refnaldi, M.Litt, and Prof. Dr. Gusril, M.Pd, as contributors who have given comments, suggestions and constructive feedback for the improvement of this thesis.

Next, it is also a pleasure to thank to the headmasters of SMP Negeri at Kec. Tambang Kab. Kampar: Zulkifli, S.Pd., M.Si. (SMPN 1), Nizar, S.Pd., M.Si. (SMPN 2), Zamira, M.Pd. (SMPN 4), Drs. M. Arianto, M.Pd. (SMPN 6), who have allowed me to collect the data in their school. Special thanks are also addressed to English teachers at SMP Negeri Kec. Tambang: Suryati, S.Pd., Amri, S.Pd., and Nelmiati, S.S., for their kindness to participate in this research.

Special words of appreciation and sincere gratitude are addressed to my beloved parents (Drs. M. Arianto, M.Pd. and Dra. Sardanis) and brothers (Iqbal, Fiqri, Farhan and Ridwan) for their continuous spirit, support, and pray so that I can finish this thesis. Then, to my special one (Rahmi Eka Putri) who supports me from the beginning until the end, and also all of my friends in UNP. The last it is hoped that this thesis will be useful for readers, especially for the improvement of the quality of semester tests in SMP Negeri at Kec. Tambang since it has an important role in the assessment of students' achievement.

Padang, Agustus 2016

The Writer

## TABLE OF CONTENTS

<b>ABSTRAK .....</b>	<b>i</b>
<b>ABSTRACT .....</b>	<b>ii</b>
<b>PERSETUJUAN AKHIR TESIS.....</b>	<b>iii</b>
<b>PERSETUJUAN KOMISI UJIAN TESIS.....</b>	<b>iv</b>
<b>SURAT PERNYATAAN .....</b>	<b>v</b>
<b>ACKNOWLEDGEMENT .....</b>	<b>vi</b>
<b>TABLE OF CONTENTS .....</b>	<b>vii</b>
<b>LIST OF TABLES .....</b>	<b>ix</b>
<b>LIST OF FIGURES .....</b>	<b>x</b>
<b>LIST OF APPENDICES .....</b>	<b>xi</b>
 <b>CHAPTER I: INTRODUCTION</b>	
A. Background of the Research.....	1
B. Identification of the Problem .....	4
C. Limitation of the Research.....	5
D. Formulation of Research Problem .....	5
E. Research Questions.....	5
F. Purposes of the Research .....	6
G. Significance of the Research.....	7
H. Definition of key Terms.....	8
 <b>CHAPTER II: REVIEW OF RELATED LITERATURE</b>	
A. Review of Related Theories.....	9
1. Definition of Test.....	9
2. Types of Test.....	10
3. Characteristics of Good Test.....	13
4. Validity .....	17
5. Reliability .....	18
6. Teachers' Factors Affecting the Quality of Tests.....	19
7. Items Analysis .....	23
B. Review of Related Previous Studies .....	27
C. Conceptual Framework .....	29
 <b>CHAPTER III: RESEARCH METHODOLOGY</b>	
A. Type of the Research .....	31
B. Source of Data .....	31
C. The Population and Sample .....	32
D. Instrumentation .....	34
	vii



E. Technique of Data Collection .....	35
F. Technique of Data Analysis .....	36

#### **CHAPTER IV: FINDINGS AND DISCUSSION**

A. Findings .....	40
1. Semester Test at SMPN 1 Tambang.....	40
2. Semester Test at SMPN 2 Tambang.....	56
3. Semester Test at SMPN 4 Tambang.....	71
4. Semester Test at SMPN 6 Tambang.....	86
B. Discussions .....	104
1. Validity .....	104
2. Reliability .....	108
3. Difficulty Indices.....	109
4. Discrimination Indices.....	110
5. Function of Distractors.....	111
6. Factors Affecting the Quality of the Test.....	112
C. Limitation of the Research.....	114

#### **CHAPTER V: CONCLUSIONS AND SUGGESTIONS**

A. Conclusions.....	115
B. Implications .....	116
C. Suggestions .....	116

<b>BIBLIOGRAPHY</b> .....	118
---------------------------	-----

<b>APPENDICES</b> .....	121
-------------------------	-----



## LISTS OF TABLES

Table 1. The Number of Population.....	32
Table 2. The Number of Sample.....	33
Table 3. Data of Score Distribution at SMPN 1 Tambang .....	41
Table 4. Basic Competences for Semester Test at SMPN 1 Tambang .....	42
Table 5. Distribution of Basic Competences in the Test Items at SMPN 1 Tambang .....	46
Table 6. The Validity of Test Items at SMPN 1 Tambang .....	48
Table 7. Difficulty Indices of the Test Items at SMPN 1 Tambang .....	51
Table 8. Discrimination Indices of the Test Items at SMPN 1 Tambang .....	53
Table 9. Function of Distractors of the Test Items at SMPN 1 Tambang.....	55
Table 10. Data of Score Distribution at SMPN 2 Tambang .....	56
Table 11. Basic Competences for Semester Test at SMPN 2 Tambang .....	58
Table 12. Distribution of Basic Competences in the Test Items at SMPN 2 Tambang .....	61
Table 13. The Validity of Test Items at SMPN 2 Tambang .....	64
Table 14. Difficulty Indices of the Test Items at SMPN 2 Tambang .....	67
Table 15. Discrimination Indices of the Test Items at SMPN 2 Tambang .....	68
Table 16. Function of Distractors of the Test Items at SMPN 2 Tambang .....	70
Table 17. Data of Score Distribution at SMPN 4 Tambang .....	71
Table 18. Basic Competence for Semester Test at SMPN 4 Tambang.....	73
Table 19. Distribution of Basic Competences in the Test Items at SMPN 4 Tambang .....	77
Table 20. The Validity of Test Items at SMPN 4 Tambang .....	80
Table 21. Difficulty Indices of the Test Items at SMPN 4 Tambang .....	83
Table 22. Discrimination Indices of the Test Items at SMPN 4 Tambang .....	84
Table 23. Function of Distractors of the Test Items at SMPN 4 Tambang .....	85
Table 24. Data of Score Distribution at SMPN 6 Tambang .....	87
Table 25. Basic Competence for Semester Test at SMPN 6 Tambang.....	88
Table 26. Distribution of Basic Competences in the Test Items at SMPN 6 Tambang .....	92
Table 27. The Validity of Test Items at SMPN 6 Tambang .....	95
Table 28. Difficulty Indices of the Test Items at SMPN 6 Tambang .....	98
Table 29. Discrimination Indices of the Test Items at SMPN 6 Tambang .....	100
Table 30. Function of Distractors of the Test Items at SMPN 6 Tambang .....	101
Table 31. The Comparison of Semester Tests Quality at SMPN Tambang .....	102



## **LISTS OF FIGURES**

Figure 1. Conceptual Framework .....	30
--------------------------------------	----



## **LISTS OF APPENDICES**

Appendix 1. Lembar Validasi Instrumen .....	121
Appendix 2. Basic Competences for English Grade VIII SMP .....	122
Appendix 3. Interview Guideline .....	126
Appendix 4. Students' Score .....	127
Appendix 5. Result of Pearson-Product Moment Correlation .....	131
Appendix 6. Result of Difficulty Indices and Discrimination Indices .....	146
Appendix 7. Result of Function of Distractors .....	157
Appendix 8. Interview Transcription .....	175
Appendix 9. Semester Tests Used at SMP Negeri Kec. Tambang .....	190
Appendix 10. Students' Answer Sheets .....	201
Appendix 11. Surat Izin Penelitian .....	209
Appendix 12. Surat Keterangan Telah Melakukan Penelitian .....	210



## **CHAPTER I**

### **INTRODUCTION**

#### **A. Background of the Problem**

There are some ways which can be used to evaluate teaching and learning activities; one of them is by giving a test. Test is an instrument to measure students' performance in certain skills or abilities. Testing is very important in teaching learning process since it aims to evaluate students' understanding towards the materials or to analyze whether the goal of the education itself has been achieved or not. According to Brown (2010: 3), a test is a method of measuring a person's ability, knowledge or performance in a given domain. Through tests, teacher can know how far students understand materials or the effectiveness of their performance in teaching. In other word, the teacher can get information about how well students have mastered the courses, which they have just learned.

As an instrument of evaluation, a test should be valid and reliable. Validity and reliability of a test are very essential since they provide the information which is important for decision making and judgment by the teachers at school. In addition, a good test will provide real information and data about student achievement and teaching practice as well as the school progress in implementing the contents of curriculum. If the test gives inaccurate information about student achievement, it can cause a wrong decision making and judgment for improving the quality of teaching program.



Dealing with the application at schools, the tests are given to measure students' understanding about their lesson. In Indonesia, one kind of tests used at school is semester test which is given to students at the end of the semester. It aims to measure what students have achieved at the end of a course. They give teachers information about how well students have accomplished learning objectives. Semester tests are usually designed by a group of teachers (MGMP) in the region. This kind of test is usually called as standardized test since it is designed by teachers who are qualified and chosen according to their educational background and experiences.

This also happened at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau. Semester tests in that area were usually designed by MKKS (Musyawarah Kerja Kepala Sekolah). This means that MKKS selected some particular teachers to design the test. The teachers have good qualification and experience in designing a good test. However, for academic year 2015/2016, the semesters test, including English subject, in most Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau had been designed by the teacher who taught at the school. It implied that the headmaster from each school asked the teacher from each subject at each grade to design the semester test. This is usually called teacher-made tests. This new regulation showed that each school had different semester tests. In this case, teacher-made tests played the central role in the evaluation of students' learning in Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau since student assessment depends on

teacher-made tests except at the end of the last term when students take the National Final Exam constructed by the government.

Teacher-made test is usually used for formative assessment. Wiggins (1989) states that teacher-made tests are usually criterion referenced tests designed to assess students' mastery of a specific body of knowledge. In other words, it can be stated that teacher-made test is actually a kind of formative tests. Teacher-made test is never used as the summative test before since it varies according to teachers' experiences and expectation in designing the test. Moreover, studies on the impact of formative test on students' achievement in summative tests have not been very positive. According to Burton and Calfee (1989) in Kinyua and Okuya (2014), the problem of using formative tests for evaluation is that teacher-made tests are often severely flawed. In addition, Wiggins (1989) argues that most criterion-referenced tests are inadequate because the problems are contrived and the cues artificial. Therefore, Parr and Bauer (2006) in Kinyua and Okuya (2014) suggest that if teacher-made tests are going to be used as summative tests, those tests must be comparable on the key attributes of the test quality such as validity and reliability.

Based on the explanation above, it can be concluded that there are several factors which give contributions to the quality of a test, such as the validity and reliability of the test. Moreover, the teachers' experience and educational background also influence the quality of the test they design. Thus, teacher-made tests should fulfill these requirements in order to make them qualified to become a good test. However, based on the observation done to several tests used at SMP



Negeri Kec. Tambang, it was found that there were some weaknesses in the tests, especially in the construction mechanism of the tests. There were several misspelled words and unclear instructions. Some questions even did not attach the text. For example, in one of the tests it was written:

25. *Who sent the letter?*

*a. Marsden Holiday*

*c. R.J. Evans and his wife*

*b. R.J. Evans*

*d. A tour operator*

This item asked students about the writer of the letter while there was no letter attached to the question. Another case found was that there were some numbers skipped. As a result, it was difficult for the students to answer the questions. Moreover, it would affect the quality of the tests itself. For this reason, the researcher intended to seek and analyze more deeply the quality of teacher-made tests for semester tests in Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau.

## **B. Identification of the Problem**

Related to the background of the problem mentioned above, it is clear that a test, as an instrument to assess students' learning achievement, has an important role in teaching and learning process. Therefore, teachers should consider the criteria of a good test to make the test have a good quality. Starting from the academic year 2015/ 2016, the schools at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau used the tests designed by the teachers as the instrument to assess students' achievement in each semester. However, several weaknesses were found the semester tests made by the teachers.

Hence, the quality of the teacher-made tests should be analyzed more deeply in order to evaluate whether the tests fulfill the criteria of good tests and can be used as good summative tests.

### **C. Limitation of the Problem**

Based on the identification of the problem above, this research was limited to the semester tests made by the Grade VIII teachers for Sekolah Menengah Pertama (SMP) at Kec. Tambang Kab. Kampar Riau. Those tests were used to evaluate students' achievement in the first semester of academic year 2015/ 2016. The researcher limited the analysis to the validity, reliability and also the item analysis of the tests made by the teachers. In addition, the researcher also wanted to find out the factors from the teachers which influence the quality of teacher-made tests.

### **D. Formulation of Research Problem**

Based on the background of problem above, this research problem was formulated as follows: *How is the teacher-made test for Grade VIII of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau?*

### **E. Research Questions**

Based on the formulation of the problem above, the researcher elaborated the main problem into the following questions:



1. How is the validity of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau?
2. How is the reliability of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau?
3. How are the difficulty indices, item discriminations, and function of distractors of the semester tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau?
4. What factors from teachers influenced the quality of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau?

#### **F. Purposes of the Research**

The purpose of this research was to investigate some issues related to the teacher-made tests for Grade VIII of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau. The purposes of this research were as follow:

1. To find out the validity of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau.

2. To find out the reliability of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau.
3. To find out the difficulty indices, item discriminations, and function of distractors of the semester tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau.
4. To investigate factors from teachers influenced the quality of the tests made by the Grade VIII teachers of Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang Kab. Kampar Riau.

### **G. Significance of the Research**

The result of this research is expected to give significant contributions theoretically and practically. Theoretically, it is expected that the result of this research can enrich the knowledge and understanding about test evaluation, especially the quality of teacher-made tests. In addition, it is also expected to be a suggestion for designing better teacher-made tests. Meanwhile, practically, the result of this research is expected to be helpful for English teachers especially at Sekolah Menengah Pertama (SMP) Negeri at Kec. Tambang, Kab. Kampar Riau and other teachers to improve the quality of the tests they design. Then, the result of this research is also expected to give suggestions to the headmaster association (MKKS) in deciding whether the semester tests are designed by the teachers in each school or by MKKS. Moreover, for other researchers who want to conduct a



research in this field, this research can be a reference about the validity, reliability and item analysis of teacher-made tests and to become a base to conduct a more comprehensive research about this topic.

## **H. Definition of Key Terms**

The key terms in this research are as follow:

1. Analysis : A careful study of something to learn about its parts, what they do, and how they are related to each other.
2. Teacher-made tests : The semester tests that are made and used by English teachers of SMP Negeri at Kec. Tambang Kab. Kampar, Riau.
3. Validity : The criteria to measure whether a test can measure accurately.
4. Reliability : The consistency of a test.
5. Item Analysis : Statistical evaluation about the effectiveness of test items.

## **CHAPTER V**

### **CONCLUSIONS AND SUGGESTIONS**

#### **A. Conclusion**

Based on the result of the research, it can be concluded that:

1. The validity of the tests is varied based on the content of the items included in tests and the item validity index of the tests. It was found that the validity of the tests are varied as a result of several factors that threaten the validity of the tests, such as unclear test directions, confusing test items, inconsistent scoring methods and untaught items included in the test.
2. The reliability of the tests is still low as it was found that the index of reliability coefficient is categorized as low; only one test can be categorized to have moderate reliability.
3. The difficulty indices of the tests are categorized as moderate. This happens because of providing clues in the stems and using key words as key answers. Meanwhile, the discrimination indices of the tests are categorized as satisfactory as it was found that most tests could discriminate between high and low students. This happens because most distractors had functioned well.
4. There is only one factor from the teachers which significantly influences the quality of the test they make, which is training on tests construction. Although the teachers have high educational level and experience in test construction, their ability in constructing a good test will not improve if they have not got the training about test constructions.



**B. Implications**

The research findings bring some practical implications to the test development. They are as follow:

1. The test should be designed by considering about the characteristics of a good test including the validity, reliability, and item analysis of the test.
2. The content area which is intended to measure should be written in the test blueprints in which the basic competences, item indicators, materials, and number of items are stated clearly. Thus, teachers of SMP Negeri at Kec. Tambang should make the test blueprints to help them design a good test.

**C. Suggestions**

Based on the research findings, the writer wants to provide some suggestions, such as:

1. English teachers should try out, analyze, and revise the tests which they design before giving to the students. It is important to ensure that the tests are reliable and valid as an instrument to measure students' achievement. Besides, the teachers should also participate more frequently in some communities concerning about test construction.
2. The Headmaster Association (MKKS) and schools at Kec. Tambang need to invite practitioners of evaluation or language testing more frequently to train the teachers about test constructions. It is because designing test is one of the teachers' duties at schools. English teachers usually evaluate students'

progress by using tests. Besides, designing a good test is not easy. Hence, well training and knowledge about test design are necessary for English teachers.

3. MGMP at Kec. Tambang should be more active to prepare and discuss tests that will be used by the teachers. This can be a good way to share knowledge and experience of making good tests.
4. Dinas Pendidikan at Kab. Kampar should provide more training on test construction.
5. It is expected that other researchers will conduct a more comprehensive research on the evaluation of test and test design.

## BIBLIOGRAPHY

- Agu, Ngozi N., Chika Onyekuba, and Aloysius C. Anyichie. 2013. Measuring Teacher' Competencies in Constructing Classroom-based Tests in Nigerian Secondary Schools: Need for a Test Construction Skill Inventory. *Academic Journals of Educational Research and Reviews*, Vol. 8 (8,) April 23, 2013. ISSN: 1996-0816.
- Arikunto, S dan Cepi Safruddin Abdul Jabar. 2014. *Evaluasi Program Pendidikan: Pedoman Teoritis Praktis Bagi Mahasiswa dan Praktisi Pendidikan (Edisi Kedua)*. Jakarta: Bum Aksara.
- Arikunto, S. 2012. *Dasar-Dasar Evaluasi Pendidikan (Edisi 2)*. Jakarta: Bumi Aksara.
- Arikunto, S. 2005. *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*. Jakarta: Bumi Aksara.
- Ary, D. L. C. Jacobs and A. Razavieh. 2002. *Introduction to Research in Education (6<sup>th</sup>ed.)* Belmont: Thomson Learning, Inc.
- Bachman, L. F. 1990. *Fundamental Consideration in Language Testing*. Oxford: Oxford University Press.
- Best, J. and J. Khan. 2003. *Research in Education (9<sup>th</sup>ed.)*. Boston: Pearson Education Company.
- Brown, H. Douglas. 2010. *Language Assessment: Principles and Classroom Practice (Second Edition)*. New York: Pearson Longman.
- Brown, J. D. and T. Hudson. 2004. *Criterion-Referenced Language Testing*. Cambridge: Cambridge University Press.
- Burton, S. J., R. R. Sudweeks, P. F. Merrill and B. Wood. 1991. *How to Prepare Better Multiple Choice Test Items: Guidelines for University Faculty*. Brigham: Brigham Young University Testing Services.
- Faleye, Bamidele A. and Oluwole R. Ayeloja. 2014. Taxonomial Analysis of Selected Teacher-Made Multiple Choice Tests in Obafemi Awolowo University, Nigeria. *Journals of Educational and Social Research*, Vol. 4, May 2014. ISSN: 2239-978X.
- Finocchiaro, M. and S. Sako. 1983. *Foreign Language Testing: A Practical Approach*. New York: Regents Publishing Company, Inc.