

The Evaluation of Distractor in English Semester Exam at Junior High School

Armilia Riza^{1}, Hevriani Sevrika², Elmiati³, Muthia Tiara⁴*

¹*Pendidikan Bahasa Inggris, Universitas PGRI Sumatera Barat*

**email: rizaridwan23@gmail.com, sevrikariecha@gmail.com, elmiatisofiana@gmail.com*

Abstract

This study aims to determine the quality of the items in the Odd Semester Final Examination Questions for English Subjects at SMP NEGERI 09 Padang for the 2022/2023 Academic Year. This research is quantitative descriptive. The Odd Semester Final Examination Items for English Subjects of SMP NEGERI 09 Padang consist of 50 multiple choice questions. This type of research is descriptive quantitative research. The participants of this study were 70 students of class VIII at SMP N 09 Padang which were determined through simple random sampling technique. This instrument is documentation of student answer sheet and key answer. From the data analysis the researcher can conclude that in terms of the effectiveness of the distractor, there are 19 items (38%) belonging to the very good and 15 items (30%) categories, 11 items (22%) including the sufficient category, 3 questions (6%) including the poor category, and 2 questions (4%) the questioners fall into the bad category

Key words: *Evaluation, Item analysis, Anates*

Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Bahasa Inggris di SMP NEGERI 09 Padang Tahun Pelajaran 2022/2023. Penelitian ini bersifat deskriptif kuantitatif. Soal Ujian Akhir Semester Gasal Mata Pelajaran Bahasa Inggris SMP NEGERI 09 Padang terdiri dari 50 soal pilihan ganda. Jenis penelitian ini adalah penelitian deskriptif kuantitatif. Partisipan penelitian ini adalah 70 siswa kelas VIII SMP N 09 Padang yang ditentukan melalui teknik simple random sampling. Instrumen ini berupa dokumentasi lembar jawaban siswa dan kunci jawaban. Dari analisis data peneliti dapat menyimpulkan bahwa dari segi keefektifan distraktor terdapat 19 butir (38%) yang termasuk kategori sangat baik dan 15 butir (30%), 11 butir (22%) termasuk kategori cukup, 3 pertanyaan (6%) termasuk kategori kurang baik, dan 2 pertanyaan (4%) penanya termasuk dalam kategori kurang baik.

Kata Kunci: *evaluasi, butir soal, anates*

INTRODUCTION

Evaluation is a planned activity to measure, assess, and the success of a program. The goal of the evaluation is to determine the level of achievement of learning objectives and to assess the efficacy of teaching and learning. A test is a reliable tool that is frequently used by teachers to assess the quantity and quality of learning completed. One of the most widely used assessment techniques in education is testing. Tests can be

viewed as both instruments and as standardized procedures that entail asking a series of questions in order to methodically measure a sample of behaviour. Tests are created to compare a sample's quality, competency, or comprehension to a requirement that is commonly viewed as acceptable or not. Tests are instruments used in educational practice to assess a student's capacity for finishing a task, proving

competence in a skill or demonstrating subject understanding. Both weekly spelling tests and multiple-choice exams are permissible (Adom et al, 2020).

According to (Danuwijaya, 2018) Item analysis is a test assessment in order to obtain a set of questions that have quality adequate for each item of the question. Analysis of the items that have been answered by students has an important purpose such as answers on the questions that have been tested and then analyzed. Item analysis is an important aspect of the test development process since it identifies things that should be changed in terms of quality for future tests or perhaps eliminated owing to misleading information. Examining students' and teachers' academic growth and achievement is the purpose of giving assessments in the classroom. Students' understanding of the course materials can be evaluated by a test, which can also be used to gauge how well teachers have accomplished the learning objectives for each unit of the lesson. Essays, multiple-choice questions, true-false questions, and matching questions are just a few of the exam topics the teacher can use to evaluate the students. The multiple-choice test item is the one that is most usually used in schools, though, as it is quick and efficient at measuring a range of knowledge, skills, attitudes, and abilities (Riza & Afriyanti, 2016)

Seeing the importance of evaluation as a means of achieving the goals of a learning process, and the importance of evaluation through item analysis based on the level of difficulty and distractor that has been described and with this gap. The teacher should collect, summarize, and use student

input when they are analyzing the subject, and they should use the test results to draw conclusions about the students' performance. Item analysis serves a number of objectives, including analyzing if the item satisfies the teacher's expectations, explaining the ability as a result, addressing the problem with students in the classroom, and providing feedback to the teacher and lecturer on the students' problems. Additional advantages include giving feedback on particular areas of curriculum development, changing the materials, and enhancing the capacity to create test items (Helmanda, 2019).

Based on the above the researcher has described, the researcher can thus identify the problem. The test is evaluated to have failed of demonstrating the distractor's effectiveness. Moreover, item analysis can be used to know whether the students do well in their exam and can find the distractor.

The researcher is interested to investigate the use one of tools to analyze the final items the teachers need is web application ANATES version 4. ANATES Version 4 is a software tool that analyzes multiple-choice and explanation questions.

RESEARCH METHOD

The researcher used descriptive research because the researcher wanted to describe about analyses of students final examination by using anates program at SMP Negeri 9 Padang. This study was carried out on the items of the VIII class English subject exam at SMP Negeri 9 Padang, to find out what the quality of the item by conducting quantitative analysis using aspects of quality determinants. The analysis was carried out using an

application computer program to analyse the items, namely ANATES version 4. The process of examining the items quantitatively was done using empirical data from the items that were collected from the tested questions. According to Sugiyono, (2016) Study quantitative research is a research in which the data obtained are in the form of: numbers or statements that are assessed and analysed statistically. This research method is called quantitative method because the research data are in the form of: numbers and analysis.

Descriptive research is defined as a research method used to describe the existing phenomena as accurately as possible (Atmowardoyo, 2018). The participants of this study were 70 students of class VIII at SMP N 09 Padang which were determined through simple random sampling technique. This instrument is documentation of student answer sheet and key answer.

FINDING AND DISCUSSION

The researcher explained the data about an analysis of student's final examination by using Anates program at junior high school.

In the method used to determine whether a distractor can carry out its function properly or not, there are several things that need to be emphasized. The thing that applies in the world of evaluation of learning outcomes is that the distractor is declared to have been able to carry out its function properly if the distractor has been chosen by at least 5% of all participants (Pratama, 2006). A distractor can be treated in 3 ways: accepted, because it's good, rejected, because it is not good, and rewritten, because it is not good Becker et al, (2015).

Distractors in multiple choice questions the possible answers are divided

into two, namely the answer key and the distractor. Distractor serves to identify test takers who have a high ability. The purpose of distractor analysis is to find out how much many students answered correctly according to the answer key and how many choose distractors. Distractors are said to be effective if selected by test participants who come from the lower group, namely students who have low abilities. And if chosen by the student highly capable, the distractor is not works fine (Hartati & Yogi, 2019).

The effectiveness of the distractor can be calculated by the following formula (Zainal Arifin, 2014):

$$IP = \frac{P}{\frac{N-B}{n-1}} \times 100\%$$

IP = distractor index

P = the number of students who chose the distractor

N = the number of students who took the test

B = the number of students who answered correctly in question

n = number of alternative answers (options)

1 = constant

Interpretation of the calculation results for each distractor in a question item can use the following criteria.

Very good	IP = 76% - 125%
Good	IP = 51% - 75% or 126% - 150%
Moderate	IP = 26% - 50% or 151% - 175%
Poor	IP = 0% - 25% or 176% - 200%
Very poor	IP = more than 200%

Quality of semester final examination item for English Subjects in class VIII at

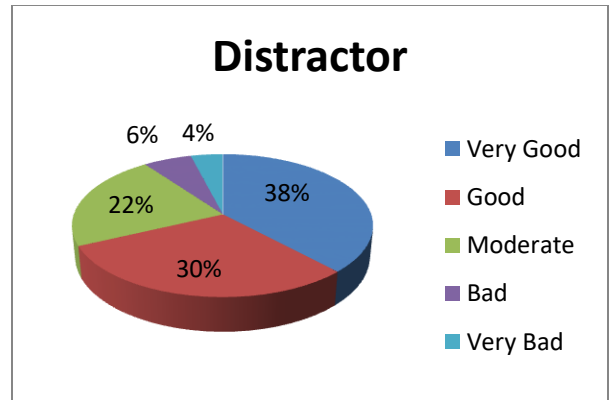
SMPN 09 Padang, viewed from the distractor of each item.

The calculation of the effectiveness of the distractor is done by using Anates program version 4. Calculation results for each distractor interpreted into five criteria, namely: IP = 76% - 125% mean very good, IP = 51% - 75% or 126% -150% means good, IP = 26% - 50% or 151% - 175% means moderate, IP = 0% - 25% or 176% - 200% means bad, and IP = more than 200% means very bad. The results showed that there were distractors who were included in the very good criteria totaling 19 items (38%) and good criteria totaling 15 items (30%), distractors who were included in the moderate criteria totaling 11 items (22%), bad amounting to 3 items (6%), very bad criteria 2 items (4%)

Tabel 1 The result of distractor

<i>Distractor</i>	<i>total</i>	<i>Items</i>	<i>Percent age</i>
Very Good	19	1,2,9,21,22,24,25,26,27,28,29,32,33,34,35,36,44,47	38%
Good	15	4,5,8,13,16,19,37,39,40,41,42,43,45,46,44	30%
Moderate	11	3,7,11,15,17,18,30,31,38,49,50	22%
Bad	3	12,14,20	6%
Very bad	2	6,10	4%

Source: primary data processed



Picture 1: Pie Chart of Item Distribution Based on the distractor

The distractor is considered good if the number of students who choose the distractor is the same or close to the ideal number. Then for the interpretation of the effectiveness of each item questions using a benchmark adapted from the Likert scale as follows.

- If the four answers that works then the item can be said to have the effectiveness of distractor is very good
- If there are three distractor answers that work then the question said to have good distractor.
- If there are two distractor answers that work then the item said to have a moderate distractor.
- If there is only one distractor answer that works then question is said to have the effectiveness of distractor is bad.
- If all the distractor answers don't work, then the question is said has very bad distractor.

Based on the results of calculations and interpretations showed that there were distractors who were included in the very good criteria totaling 19 items that are categorized as very good in numbers 1,2,9,21,22,24,25,26,27,28,29,32,33,34,35,36,44,47 (38%) and good criteria totaling 15 items are good category are found in numbers 4,5,8,13,16,19,37,39,40,41,42,43,45,46,44 (30%), distractors who were included in the moderate criteria totaling 11 items are found

in numbers 3,7,11,15,17,18,30,31,38,49,50 (22%), bad amounting to 3 items are found in number 12,14,20 (6%), very bad category 2 items are 6,10 (4%). Distractors that are bad and very bad indicate that the distractor is too conspicuous and heterogeneous, so that the distractor has no appeal for choosing tests that do not understand the material. Distractor which is bad and very bad shows that the distractor it's too conspicuous and heterogeneous, for the distractor to not have the appeal to be selected by test takers who do not understand learning material.

CONCLUSION

The researcher gives conclusion about students' final examination item analysis at junior high school. Based on the analysis of the end of first semester exam questions on English subject at SMP Negeri 09 Padang for the 2022/2023 academic year the results obtained that all questions have good criteria based on distractors, it can be concluded that: The effectiveness of the distractors on the items that were categorized as very good reached 38% (19 items), questions that were categorized as good reached 30% (15 items), questions that were categorized as moderate reached 22% (11 items), and questions that were categorized as bad reached 6% (3 items) and very bad reached 4% have (2 items).

SUGESTION

The writers suggested that the teachers who had given multiple choice test to students to try doing evaluation of test items. Thus, the teachers can be a good teacher, because the result of evaluation can measure their ability in teaching and designing the test.

ACKNOWLEDGEMENTS

Thanks to all of the students and teacher at Junior High School N. 9. Padang who have

participated on this research and also big thanks to Lecturer of English Department of Universitas PGRI Sumatera Barat for supporting this research.

REFERENCES (Examples)

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109–119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Atmowardoyo, H. (2018). Research Methods in TEFL Studies: Descriptive Research, Case Study, Error Analysis, and R & D. *Journal of Language Teaching and Research*, 9(1), 197. <https://doi.org/10.17507/jltr.0901.25>
- Becker, F. G., Cleary, M., Team, R. M., Holtermann, H., (2015). Analysis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Syria Studies*, 7(1), 37–72.
- Danuwijaya, A. A. (2018). Item Analysis of Reading Comprehension Test for Post-Graduate Students. *English Review: Journal of English Education*, 7(1), 29. <https://doi.org/10.25134/erjee.v7i1.1493>
- Helmanda, R. N. & C. M. (2019). *The Classification of Reading Comprehension Test Items*.
- Pratama, I. nyoman D. (2006). Evaluasi pendidikan. *Jakarta: Rineka Cipta*, 267–294.
- Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (p. 7).