



## **AN ANALYSIS OF TEST ITEMS IN ENGLISH SUBJECT FINAL TEST OF TEN GRADE STUDENTS AT SMA NEGERI 2 KUPANG**

**Gracela Rame, Abdul Kadir Kesi**

English Education Study Program Teacher Training and Education Faculty Artha Wacana  
Christian University, Master of Linguistics Study Program, Universitas Warmadewa  
Email: [Gracelaram2000@gmail.com](mailto:Gracelaram2000@gmail.com) , [abdulkadirkesi@gmail.com](mailto:abdulkadirkesi@gmail.com)

### **Abstract**

"An Analysis of Test Item in English Subject Final Test of Ten Grade Students at SMA Negeri 2 Kupang" is the title of this study. The following are the research's main issues: 1. How challenging is the English Subject Final Test for Tenth Grade Students at SMA Negeri 2 Kupang? 2. How strong is the English Subject Final Test Discrimination Power of Ten Grade Students at SMA Negeri 2 Kupang? 3. How successful is the English Subject Final Test Distractor for Tenth Grade Students at SMA Negeri 2 Kupang? The author employs a quantitative descriptive method to address this query. Documentation technique is the method of data collection employed. The study's findings regarding the English final exam at SMA Negeri 2 Kupang are as follows: a) There are 3 items (7.5%) that are extremely difficult, 2 items (5%) that are difficult, 2 items (5%) that are moderate, 16 items (40%) that are easy, and 17 items (42.5%) that are very easy. b) There are 1 item (2.5%) extremely poor, 28 items (70%) poor, 8 items (20%) moderate, 3 items (7.5%) good, and 0% very good in terms of distinguishing power. c) Four items (10%) are very good, eleven items (27.5%) good, thirteen items (32.5%) moderate, nine items (22.5%) poor, and three items (7.5%) very terrible in terms of distractor effectiveness. Consequently, it can be said that the questions on the Language final exam at SMA Negeri 2 Kupang were satisfactory.

**Keywords:** Item analysis, Final Test, discrimination power, the effectiveness of distractor.

### **INTRODUCTION**

One element determining the quality of human resources is education. Education will shape us into capable and valuable individuals who will progress this nation. Teachers must participate in the field of education. Teachers have enormous tasks and responsibilities as professionals in the field. Assessment is one of the teacher's responsibilities. Teachers play the function of evaluators by gathering data or information regarding successful learning initiatives. Evaluation is a methodical, long-lasting, comprehensive process that determines the value and significance of various learning components by predetermined standards and factors (Farida 2017: 2). According to Arifin (2014: 118), a test is a methodology or approach used to conduct out measuring activities, where students are required to perform or answer a series of tasks, statements, or questions to measure various elements of their behavior.

Through evaluation, teachers in language classrooms also make decisions and judgments about their pedagogical approaches. Teachers can utilize a variety of evaluation techniques in the classroom, and giving tests is one of them (Hughes, 2003). According to Arifin (2014: 118), a test is a methodology or approach used to conduct out-measuring activities, where students

are required to perform or answer a series of tasks, statements, or questions to measure various elements of their behavior. Furthermore, Gampper (2013:74) states that the exam ought to be regularly scheduled, usually at a predetermined time, and given at appropriate curriculum points. According to Osman (2010:53), the test's objectivity, discrimination, comprehensiveness, validity, reliability, specification of administering conditions, and g) dependability are the variables that determine its quality.

This suggests a relationship between the content that is taught in classrooms and the test design. You must guarantee the test's quality as a teacher or test creator. Item analysis is one technique to make sure an exam is of high quality. Item analysis, according to Sudjana (2014: 135–136), is the process of looking over test questions to create a set of questions that are sufficiently good. According to Ngatman and Fitria Andriyani (2017:1), a test is a tool used to collect information about individuals or objects. Written inquiries, interviews, checklists, or anecdotal notes that record behavior and physical characteristics can all be used as instruments. Arikunto, in the meantime (2012: 220). The methodical process of analyzing the question items will yield information that is highly relevant to the test items that we design.

## RESEARCH METHODS

This study uses quantitative descriptive research methods. Research that is conducted using the principles of objectivity and beginning with data collection, analysis, and presentation in the form of images, tables, graphs, and other visual aids is known as quantitative research. Using Microsoft Excel 2021 as a tool to analyze inquiry items, the research aims to assess how well class X teachers can design inquiries for their students. The descriptive method is the one that is applied. Researchers carry out the roles of planners, implementers, collectors, analysts, and reporters of study findings in descriptive quantitative research. The final semester test assessments of the class x IPA 4 students at SMA Negeri 2 Kupang will be used to gather the data.

Documentation is the method utilized to acquire data. The documentation method takes the form of student answer sheets and answer keys for a series of final semester exam assessment questions for class X IPA 4 students at SMA Negeri 2 Kupang. Responses from every pupil in class X, IPA 4. This study used a quantitative analytic technique for its analysis. With a dichotomy of 1 for correct scores and 0 for incorrect scores. Researchers utilize Microsoft Excel 201 to analyze data and assess the effectiveness, power of distinction, and degree of difficulty of distractions.

## RESULTS AND DISCUSSION

### 1. Difficulty level

The researchers used the Arikunto formula to account for item difficulty in order to determine the level of difficulty. One way to accomplish this is to figure out what percentage of test participants provide accurate answers.

$$D = \frac{P}{J_s}$$

Where :

D = difficulty index.

P = The number of students who answer the questions correctly

J<sub>s</sub> = The total number of students participating in the test

The difficulty level scale spans from 0.00 to 1.00. It is then ranked into a number of categories. This classification system interprets the degree of difficulty. The detail rank is shown in the table below:

Table 1. criteria Difficulty Level

| Descrimination Power | Interpretion     |
|----------------------|------------------|
| P 0,00               | Verry difficulty |
| 0,00 < P ≤ 0,30      | Difficulty       |
| 0,30 < P ≤ 0,70      | Desirable        |
| 0,70 < P ≤ 1,00      | Easy             |
| P ≤ 1,00             | Verry easy       |

(Arikunto, 2005:210)

Table 2. The Result item difficulty

| No | Category        | Items   | Total | Present |
|----|-----------------|---|-------|---------|
| 1  | Verry Difficult | 8, 26,38  | 3     | 7,5%    |
| 2  | Difficult       | 6,7,  | 2     | 5%      |
| 3  | Desirable       | 9,34  | 2     | 5%      |
| 4  | Easy            | 1,2,4,10,11,1<br>2,20,25,28,3<br>1<br>32,33,36,37,<br>39,40     | 16    | 40%     |
| 5  | Verry Easy      | 3,5,13,14,15,<br>16,17,18,19,<br>21<br>22,23,24,27,<br>29,30,35 | 17    | 42,5%   |

The final exam questions for the English Subject Final Test of Ten Grade Students at SMA Negeri 2 Kupang yielded results on the difficulty of the questions based on the data above. Of these, 3 or (7,5%) were very difficult questions, 2 or (5%) were difficult questions, 2 (5%) were desirable, 16 (40%) questions were Easy and 17 (42,5%) very easy.

2. Discrimination power

The researchers used the Sabri formula to account for item difficulty to determine the discrimination power.

$$D = \frac{U - L}{N}$$

D: Discrimination index

U: number of students who answer the item correctly

L: number of students who answer the item correctly

N: Total of students who participating the test

(Sabri, 2013:4)

Five levels make up the level of discrimination index, as shown in the accompanying table.

Table 3 Criteria Discrimination Power

| Discrimination Power  | Interpretation |
|-----------------------|----------------|
| $DP \leq 0,00$        | Very Poor      |
| $0,00 < DP \leq 0,20$ | Poor           |
| $0,21 < DP \leq 0,40$ | Satisfactory   |
| $0,41 < DP \leq 0,70$ | Good           |
| $0,71 < DP \leq 1,00$ | Very Good      |

(Widiyoko 2014:137)

The following table provides a summary of the findings from the calculation of the discriminant index:

Table 4 The Result of Discrimination Power

| No | Category     | Item   | Total | Present |
|----|--------------|--|-------|---------|
| 1  | Very Poor    | 38   | 1     | 2,5%    |
| 2  | Poor         | 1,2,3,4,5,8,10,13,14,15,16,17,18,19<br>20,21,22,23,24,25,26,27,28,29,30,31,35,37 | 28    | 70%     |
| 3  | Satisfactory | 7,11,12,32,33,36,39,40   | 8     | 20%     |
| 4  | Good         | 6,9,34   | 3     | 7,5%    |
| 5  | Very Good    | 0  |       |         |

Based on the data above, the results for the difficulty of the questions are obtained English Subject Final Test of Ten Grade Students at SMA Negeri 2 Kupang there are 1(2,5%) very poor, 28(70%) poor, 8(21%) satisfactory, 3(7,5%) good, 0 very Poor.

### 3. The effectiveness of distractor

The researchers used the formula to account for item difficulty to determine the discrimination power.

$$IP = \frac{P}{(N - B)/(n - 1)} \times 100$$

IP = Distractor Efficiency index

P = the number of students who choose deception

N = number of students taking the test

B = The number of students who answered correctly on each question

n = number of alternative answers (option)

( Arifin 2014:279

The level of discrimination index can be categorized into five different levels as the following table:

Table Criteria Distractor efficiency

| IP Value               | Interpretation |
|------------------------|----------------|
| 76% -125%              | Very good      |
| 51% -75% Or 126% -150% | Good           |
| 26% -50% Or 151% -175% | Average        |
| 0% -25% Or 176% -200%  | Poor           |
| More than 200%         | Very poor      |

( Arifin 2014:279)

Table 4.3 The Result of Efficiency Distractor

| Category   | Item                                 | Total | Present |
|------------|--------------------------------------|-------|---------|
| Verry Good | 7,8,9,33                             | 4     | 10%     |
| Good       | 2,4,11,12,25,26,28,31,34,38,39       | 11    | 27,5%   |
| Average    | 1,6,10,16,17,18,20,24,27,32,35,36,37 | 13    | 32,5%   |
| Poor       | 13,14,15,19,21,22,23,29,40           | 9     | 22,5%   |
| Verry Poor | 3,5,30                               | 3     | 7,5%    |

Based on the data above, the results for the difficulty of the questions are obtained English Subject Final Test of Ten Grade Students at SMA Negeri 2 Kupang there are 4 or (10 %) very Good, 11 or (27,5%) Good, 13 or (32,5%) Average, 9or (22,5%) Poor and 3 or(7,5%) very poor.

The quality of the interest questions is then ascertained by analyzing the gathered tests' level of difficulty, discrimination power, and distractor efficiency. The following criteria determine what questions are good, good enough, and bad:

1. A question is considered to be of excellent quality if it satisfies all three criteria: deception, power of distinction, and difficulty level.
2. The question is of a respectable caliber even if it only satisfies two of the three criteria.
3. An item query is deemed to be of low quality if two or more standards are not satisfied.

## CONCLUSION

Based on the results of the data analysis, the following categories can be applied to the items on the English final school test at SMA Negeri 2 Kupang for the 2018–19 academic year:

1. This is sufficient, according to the item difficulty of the final English exam items at SMA Negeri 2 Kupang in the 2018–19 academic year.
2. Based on the item discrimination power of the final English test items at SMA Negeri 2 Kupang in the academic year 2018–19, it is determined that this is sufficient.
3. Based on the Distractor Efficiency di item from the SMA Negeri 2 Kupang final English test in the 2018–19 academic year, it is deemed to be good.

## Suggestion:

We believe that the following recommendations will be helpful to educators, test publishers, and other researchers.

1. Regarding educators.

For teachers to create exam questions that are a suitable level of difficulty, it is intended that they will take students' talents into account. To ensure that the test items function as intended and don't provide unfavorable outcomes like negatives, teachers need also exercise caution when designing their exams. Worth in the ability to discriminate. To lessen subjectivity when assigning evaluations, particularly for essay questions, teachers should develop clear assessment procedures or strong assessment criteria.

2. Regarding the test publisher

It is recommended that test publishers create test items that accurately reflect students' prior knowledge and are grounded in the curriculum. Better test

items should be created by test publishers, particularly if they are to be utilized as large-scale, standardized achievement exams. Before publishing the test items that are used to assess the pupils, test publishers ought to carry out some study or a trial run.

3. Regarding further researchers

Before conducting additional research that is linked to this topic, other researchers can utilize the findings of this study as a guide or point of reference. It would be preferable if additional researchers looked into this study more thoroughly or conducted an analysis.

## REFERENCES

- Arifin, Z. (2014). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arikunto, S. (1999). *Dasar-Dasar Evaluasi Pendidikan Edisi Revisi*. Bumi Aksara.
- Farida, I. (2017). *Evaluasi Pembelajaran Berdasarkan Kurikulum Nasional (2nd ed.)*. PT. Remaja Rosdakarya.
- Gampper, C. (2013). Improving English Test Qualities. *Thammasat Review*, 16(3).
- Ngatman, & Andriyani, F. D. (2017). *Tes dan Pengukuran untuk Evaluasi dalam Pendidikan Jasmani dan Olahraga*. (M. Fadilah (ed.); 1st ed.).
- Osman, R. M. (2010). *Educational Evaluation and Testing*. African Virtual University Press.
- Sabri, S. (2013). Item Analysis of Student Comprehensive Test for Research in Teaching Beginner String Ensemble Using Model Based Teaching Among Music Students in Public Universities. *International Journal of Education and Research*, 1(12).
- Sudjana, N. (2014). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya
- Sukardi. (2003). *Metode Penelitian Pendidikan dan Pengembangannya (First Edition)*. Jakarta: Bumi Aksara.
- Widoyoko, E. P. (2014). *Penilaian Hasil Pembelajaran di Sekolah (Cetakan Pertama)*. Yogyakarta: Pustaka Pelajar.