THE ANALYSIS OF ITEM DIFFICULTY AND DISCRIMINATION OF ENGLISH SUMMATIVE TEACHER MADE TEST (TMT) AT MAS MADANI PAO-PAO GOWA

Muhammad Syahruddin Nawir
Universitas Islam Negeri Alauddin Makassar
muhammad.syahruddin.nawir@uin-alauddin.ac.id

Serliah Nur
Universitas Islam Negeri Alauddin Makassar
serliah.nur@uin-alauddin.ac.id

Multazam Abubakar
Universitas Islam Negeri Alauddin Makassar
multazam.abubakar@uin-alauddin.ac.id

Andi Asmawati
Universitas Islam Negeri Alauddin Makassar
andi.asmawati@uin-alauddin.ac.id

Maemunah
Universitas Islam Negeri Alauddin Makassar
maemunah@gmail.com

ABSTRACT

This research was aimed to find out the item difficulty and item discrimination power of English summative test for the second grade students at MAS Madani Pao-Pao Gowa. The researcher used descriptive quantitative method to desribe the item difficulty and discrimination of the items of English summative test. English summative final test used to test students registered as second-grade students at MAS Madani Pao-Pao Gowa in the academic years 2022-2023 was the subject of this study. The instruments documents in the form of test with 25 multiple-choice questions, student's answers, and an answer key. The result of this research, item difficulty is the easy category, there are 2 difficult items, 7 medium items, and 16 easy items. Meanwhile, item discrimination is the poor category, there are 0 excellent items, 0 good items, 0 sufficient items, 10 insufficient items, 15 poor items. Item difficulty and item discrimination are not equal so that it needs to be considered by the teacher when making questions. The items arranged are in accordance with the assessment for students' understanding and abilities.

Keywords: Discrimination, Item Analysis, Item Difficulty, Summative Test, Teacher-Made Test (TMT)

INTRODUCTION

Evaluation is important in teaching and learning in order to investigate the outcomes of the process. Evaluation is connected with the teaching and learning process and cannot be separated. In the formal education system, a teacher is one of the charges of the success or failure of the learning process. A teacher must assess student progress during the teaching process in terms of their understanding of the lessons they have been taught over a specific amount of time. Evaluation in education can be viewed as both a formal and informal method of assessing

students' academic achievement. Informal evaluation is commonly used during the teaching and learning process. Teachers can assess their students' achievement by witnessing and passing judgment on their student's performance during the teaching and learning process.

Evaluation is conducted to increase the capability of an organization. In other words, by doing an evaluation the teachers to measure if the learning and teaching activity is effective as it has been planned. Djiwadono (2011) describes that four classification of evaluation test based on educational test enforcement, such as formative, summative, pretest, and posttest. Summative tests provide a criterion of achievement for students after a long period of treatment by a teacher from a certain subject, in this example by the English teacher. Summative tests are evaluation of student learning and teacher teaching after learning is carried out (Brookhart & Nitko, 2018). The best can also be used to designate course grades in order to indicate students' achievement of targeted learning outcomes (Sengkaton et al., 2020). According to the above explanation, the summative test is given sometimes to ensure the students' comprehension of the materials.

To evaluate students' achievement in the subject that has been given, the teacher can usually give them some questions in the form of tests. Teachers can administer the accomplishment test at the end of each semester. Formative testing and summative tests are the two forms of accomplishment tests. The researcher chose a summative test for analysis in this study. The difficulty of the problem is how easy and how difficult a problem is for students (Kocdar et al., 2016). The percentage of students who successfully answered the questions represents the level of difficulty. The discrimination can be calculated based on the division of groups, namely the upper group is classified as intelligent, and the lower group is a group of students who are classified as not intelligent. If there is a discriminating power index number with a positive sign, it is possible to say that the thing has discriminating power, more participants in the intelligent category can answer correctly, while participants who were included in the category of not clever more who answered incorrectly (Fadlilah, 2018). Therefore, the difficulty to know the easy problem or not, but the discrimination to know which students answered more correctly and more incorrectly. Several problems students have before joined the test, students don't have quotas or the network is unstable, and before submitting they forget to check back questions that have not been answered.

In this research, the researcher used a summative test, which is a sort of test given after the semester to assess what students achieved both individually and in groups. The test can be an essay or multiple-choice test. The teacher who makes the exam must know the steps to take in making a good exam. A test is said to be a good test if it meets some of the conditions of a good test. The validity and reliability of the questions is one technique to determine the quality of a test.

There are several reasons the teacher-made English summative test for second-grade students MAS Madani Pao-Pao Gowa was chosen. First, the teacher needs to plan a good test. Teachers or researchers have to develop an effective test for the results to be valid and reliable. In terms of a

good test, Mardapi in Maharani & Putro (2020) stated there are nine steps to develop a highly qualified test, they are making a test, developing a test, analyzing a test, conducting an evaluation, analyzing test items, correcting the test, constructing the test, administering the test, and reviewing the test findings. Applying this procedure can help teachers and test makers develop a well-designed test. Second, the researcher has analyzed the summative problem of the English teacher-made test (TMT). Third, the researcher assessed students' comprehension of the subject provided. Thus, this research aims at analyzing the item difficulty of English summative items teacher-made test (TMT) and investigating the item discrimination of English summative teacher made test (TMT) for the second-grade students of MAS Madani Pao-Pao Gowa.

METHOD

Design

This research is a descriptive quantitative method. The researcher analyzed item difficulty and item discrimination of English summative teacher-made tests. In analyzing the content item difficulty and discrimination, the researcher described the conformity and the unconformity of the English teacher-made test (TMT) in MAS Madani Pao-Pao Gowa. The researcher has performed calculations to determine the correct answers for each item, the total correct answers, and the value of each item, tabulating the results into a table, total variance, and testing the test's item difficulty and discrimination.

Instrument

The instruments of research are documents in the form of teacher-made test (TMT), the student's answer, and the answer key. The teacher-made test consisting of 25 multiple-choice question numbers, the answer sheet contained the students' test responses which consists 39 students, which were used to examine the item difficulty and discrimination of the test, whilst the answer key served as a reference for knowing and determining students' scores.

Data Analysis

1. Counts the number of correct and incorrect questions.
2. Do the same process for each item question.
3. Analysis of each item of English summative tests.
4. The formulas for finding understanding are adapted from Boopathiraj (2013) as follows:

$$\mathbf{IF} = \frac{\boldsymbol{UG + LG}}{\mathbf{N}}$$

IF = Item of facility

UG = The number of correct answers provided by the upper group

LG = The number of correct answers provided by the lower group

N = The number of students who completed the test

Table 1. The Classification of Difficulty Index

No	Range	Category
1	0.00 – 0.30	Difficult
2	0.30 – 0.70	Medium
3	0.70 – 1.00	Easy

5. The second formula, to find the discrimination item is adapted from Boopathiraj (2013) as follows:

6. $$\mathbf{ID} = \frac{UG - LG}{N}$$

ID = Item discrimination

N = The number of students in one group (1/2 N)

UG = Score frequency by upper group (upper half)

LG = Score frequency by lower group (lower half)

Table 2. The Classification of Item Discrimination Index

No	Range (%)	Category
1	84-100 %	Excellent
2	68-83 %	Good
3	52-67 %	Sufficient/enough
4	36 – 51 %	Insufficient
5	0-35 %	Poor

FINDINGS AND DISCUSSION

Findings

Item Difficulty

The difficulty level of an item simply indicates how easy or tough the item performed in the exam. The difficulty level is typically stated as a fraction of the percentage of students that correctly answered the item. Although it is practical to right away take an item's complete correct answer. The research intends to compare students who performed well on the whole test and students who performed poorly on the whole test. The item difficulty of English summative teacher made test (TMT) for the second-grade students of MAS Madani Pao-Pao Gowa.

Table 3. The Data Presentation of Test Item Difficulties

Item	UG	LG	N	Category
1	14	8	38	Medium
2	17	10	38	Medium
3	18	17	38	Easy
4	18	14	38	Easy
5	19	13	38	Easy
6	17	11	38	Easy
7	19	11	38	Easy
8	17	13	38	Easy
9	2	0	38	Difficult
10	17	8	38	Medium
11	15	8	38	Medium
12	15	12	38	Medium
13	19	17	38	Easy
14	8	2	38	Difficult
15	19	16	38	Easy

16	19	12	38	Easy
17	17	17	38	Easy
18	16	12	38	Easy
19	18	12	38	Easy
20	18	9	38	Medium
21	10	5	38	Medium
22	19	11	38	Easy
23	19	18	38	Easy
24	18	16	38	Easy
25	19	18	38	Easy

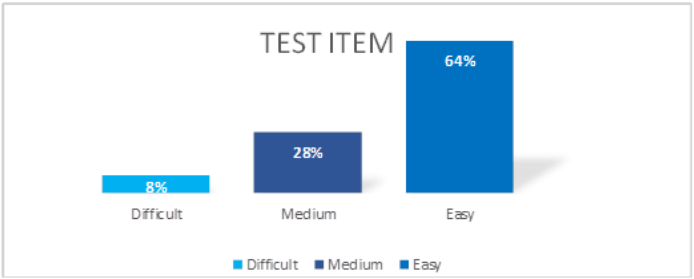
From Table 3 above, it could be seen that the classification of the difficulty index of the test items was: 2 items are difficult, 7 items are medium, and 16 items are easy from 38 students.

Table 4. Category Classification of UG and LG

Item	UG	LG	Category
1	14	8	Medium
2	17	10	Medium
3	18	17	Easy
4	18	14	Easy
5	19	13	Easy
6	17	11	Easy
7	19	11	Easy
8	17	13	Easy
9	2	0	Difficult
10	17	8	Medium
11	15	8	Medium
12	15	12	Medium
13	19	17	Easy
14	8	2	Difficult
15	19	16	Easy
16	19	12	Easy
17	17	17	Easy
18	16	12	Easy
19	18	12	Easy
20	18	9	Medium
21	10	5	Medium
22	19	11	Easy
23	19	18	Easy
24	18	16	Easy
25	19	18	Easy

Based on the table 4.2 provide that the difficult items, they are number 9, it consist of UG 2 students could answer item and LG all the students could not answer item, but in number 14, it consist of UG 8 students could answer item and LG 2 students could answer item. While number item 21 in LG there are 5 students could answer item, it is category medium. Number nine only two students can answer it. In this case means that item is very difficult.

Figure 1. Category Classification of Items



Based on the figure above showed that classification of items consisted of three categories, they are items 2 items (8%) are difficult, 7 items (28%) are medium, 16 items (64%) are easy.

Table 5. Classification Category of Difficult, Medium, and Easy

Item	UG	LG	Category
1	14	8	Medium
2	17	10	Medium
3	18	17	Easy
4	18	14	Easy
5	19	13	Easy
6	17	11	Easy
7	19	11	Easy
8	17	13	Easy
9	2	0	Difficult
10	17	8	Medium
11	15	8	Medium
12	15	12	Medium
13	19	17	Easy
14	8	2	Difficult
15	19	16	Easy
16	19	12	Easy
17	17	17	Easy
18	16	12	Easy
19	18	12	Easy
20	18	9	Medium
21	10	5	Medium
22	19	11	Easy
23	19	18	Easy
24	18	16	Easy
25	19	18	Easy

As shown in table 4.3 above, the classification of the difficulty index of the exam items was:

1. Two items are difficult = 8%
2. Seven items are medium = 28%
3. Sixteen items are easy = 64%

Item Discrimination

Item discrimination is a statistic that illustrates the extent to which an item differentiates students who perform well from students who perform poorly in the overall test. Item discrimination is the ability of an item to discriminate between respondents who score high vs. respondents who score low.

Table 6. Discrimination Index of Test Items Data Presentation

Item	UG	LG	N	Category
1	14	8	19	Poor
2	17	10	19	Insufficient
3	18	17	19	Insufficient
4	18	14	19	Poor
5	19	13	19	Poor
6	17	11	19	Poor

7	19	11	19	Poor
8	17	13	19	Poor
9	2	0	19	Poor
10	17	8	19	Insufficient
11	15	8	19	Insufficient
12	15	12	19	Poor
13	19	17	19	Poor
14	8	2	19	Poor
15	19	16	19	Poor
16	19	12	19	Poor
17	17	17	19	Poor
18	16	12	19	Poor
19	18	12	19	Poor
20	18	9	19	Insufficient
21	10	5	19	Poor
22	19	11	19	Insufficient
23	19	18	19	Insufficient
24	18	16	19	Poor
25	19	18	19	Insufficient

Table 6 above showed that nobody could answer at number nine-item (lower group) but only two students could answer it (upper group). It meant that this item was very difficult. Besides that, 17 students could answer on number items 3, 13, and 17, and also 18 students could answer on number items 23, and 25, they are category insufficient.

Based on table 6 above, nobody could answer at number nine-item (lower group) but only two students could answer it (upper group). It meant that this item was very difficult. Besides that, 17 students could answer on number items 3, 13, and 17, and also 18 students could answer on number items 23, and 25, they are category insufficient.

Table 7. The Classification of Discrimination

N0	Test Items	Category	Total	%
1	-	Excellent	-	-
2	-	Good	-	-
3	-	Sufficient	-	-
4	2, 3, 7, 10,11, 16, 20, 22, 23, and 25	Insufficient	10	**40**
5	1, 4, 5, 6, 8, 9, 12, 13, 14, 15, 17, 8, 19, 21 and 24	Poor	15	**60**

Based on table 7 above, there was no number of the test items excellent at discriminating the test, there was no number of the test items were good at discriminating the test, there was no number of the test items was sufficient to discriminate the test, item numbers 2, 3, 7, 10,11, 16, 20, 22, 23, and 25 were insufficient and needed to be reviewed (40%) or insufficient category while item number 1, 4, 5, 6, 8, 9, 12, 13, 14, 15, 17, 8, 19, 21 and 24 were poor and needed to discarded (60%) or poor category.

Table 8. The Index of Difficulty Classification

No	Number Items	Percentage	Category
1	9, 14	8%	Difficult
2	1, 2, 10, 11, 12, 20, 21	28%	Medium
3	3, 4, 5, 6, 7, 8, 13, 15, 16, 17, 18, 19, 22, 23, 23, 24, 25	64%	Easy

Table 8 above showed that from 25 number item, it consist of three categories, they are: 8 % is difficult category, 28 % is medium category, and 64% is easy category.

Table 9. The Classification of Discrimination

No	Number Items	Percentage	Category
1	0	0	Excellent
2	0	0	Good
3	0	0	Sufficient
4	2, 3, 7, 10,11, 16, 20, 22, 23, and 25	40%	Insufficient
5	1, 4, 5, 6, 8, 9, 12, 13, 14, 15, 17, 8, 19, 21 and 24	60%	Poor

Based on the table 9 above, there was no number of the test items were excelled, good and sufficient to discrimination the test 0%, insufficient needed to reviewed (40%) or insufficient category while item number were poor and needed to discarded (60%) or poor category.

Table 10. The Data Presentation of Perception

Item	LG	N	Category	Interpretation
1	8	19	Poor	Need to discard
2	10	19	Insufficient	Need to review
3	17	19	Insufficient	Need to review
4	14	19	Poor	Need to discard
5	13	19	Poor	Need to discard
6	11	19	Poor	Need to discard
7	11	19	Poor	Need to discard
8	13	19	Poor	Need to discard
9	0	19	Poor	Need to discard
10	8	19	Insufficient	Need to review
11	8	19	Insufficient	Need to review
12	12	19	Poor	Need to discard
13	17	19	Poor	Need to discard
14	2	19	Poor	Need to discard
15	16	19	Poor	Need to discard
16	12	19	Poor	Need to discard
17	17	19	Poor	Need to discard
18	12	19	Poor	Need to discard
19	12	19	Poor	Need to discard
20	9	19	Insufficient	Need to review
21	5	19	Poor	Need to discard
22	11	19	Insufficient	Need to review
23	18	19	Insufficient	Need to review
24	16	19	Poor	Need to discard
25	18	19	Insufficient	Need to review

Table 10 above showed that from 25 number item, there is 7 item or 28% need to discard, while 18 item or 72% is need to review.

Table 11. The Facility and Discrimination Indices

Item	UG	LG	IF	ID	Remark/category
1	14	8	0.6	0.31	Improper
2	17	10	0.7	0.36	Proper
3	18	17	0.9	0.05	Improper
4	18	14	0.8	0.21	Improper
5	19	13	0.8	0.31	Improper
6	17	11	0.7	0.31	Improper
7	19	11	0.8	0.42	Proper
8	17	13	0.8	0.21	Improper
9	2	0	0.1	0.10	Improper
10	17	8	0.7	0.47	Proper
11	15	8	0.6	0.36	Proper

12	15	12	0.7	0.15	Improper
13	19	17	0.9	0.10	Improper
14	8	2	0.3	0.31	Improper
15	19	16	0.9	0.15	Improper
16	19	12	0.8	0.36	Proper
17	17	17	0.9	0.0	Improper
18	16	12	0.7	0.21	Improper
19	18	12	0.8	0.31	Improper
20	18	9	0.7	0.47	Proper
21	10	5	0.4	0.26	Improper
22	19	11	0.8	0.42	Proper
23	19	18	0.10	0.05	Improper
24	18	16	0.9	0.10	Improper
25	19	18	0.10	0.05	Improper

Based on the table 4.10 above showed that from 25 number item, there is 7 item or 28% proper category, while 18 item or 72% is improper category.

Discussion

The item difficulty is how easy and how difficult a problem is for students. The level of difficulty is indicated as the percentage of students who correctly answered the questions. Item difficulty is related to the percentage of students who effectively respond to a specific item. The level of difficulty can be determined by evaluating students' responses. It indicates that the difficulty of the questions was determined by students' responses rather than teachers perceptions. Item difficulty is a statistic that examines the percentage of students who effectively answer a specific item.

The items discrimination index provides its ability to the item discriminate against students, recognizing the more intelligent from the less intellectual. The discrimination index indicates whether students who proved well on all the questions also performed well or poorly on each item on the test. It is predicted that the overall score on the test is a valid indicator of the student's ability, namely good students performing well and students performing poorly. A good test can distinguish students' abilities. In Hotiu (2006), Item discrimination is a useful indicator of item quality when the purpose of a test is to provide a range of scores reflecting variances in student achievement, so that respondents' performances may be distinguished. It assesses the degree to which item responses distinguish between persons with a higher overall score on a test and those with a lower overall score. Boopathiraj & Chellamani (2013) explained that item discrimination of a test item is to what degree success or failure on a test item reflects possession of the ability being tested. It determines how well a given item discriminates between examinees in the item's measured function or capacity. The higher the value, the stronger the discrimination of the item. According to Brown in Baeti (2020), item discrimination is a statistic that illustrates how far an item divides students who perform well from students who perform poorly on the overall test. Item discrimination refers to an item's ability to differentiate amongst students based on how well they know the topic being assessed. To compare item replies to total test results, various hand

calculation procedures have traditionally been utilized with groups of students with upper group and lower group.

Based on the dialog with the English teacher that the students found difficulties when they took the summative test, they are: they did not have quota data, signal interference/ interruption, broken hand phone, slow loading, and late uploading when they had done/finished their summative test. Therefore, their teachers always prepare solutions as such to solve their students, they lent their hand phones/smartphones to the students. So that they could finish their summative test.

The result of the summative test, that twenty-three (23) students got to exceed, there are five (5) students got to achieve (minimum completeness criteria = 72), and eleven (11) students did not achieve or under the minimum completeness criteria. Of 11 students who did not achieve minimum completeness criteria (MCC), 7 students found difficulties when they had a summative test and the others did not find difficulties there.

Based on the findings that the ability of the students in MAS Madani Pao-Pao Gowa who are in the upper group (UG) and lower group (LG) is not the same. The findings proved that there are significant differences in the ability of students in the upper group and lower group to answer test items. In addition, this indicates that an item can distinguish between upper-group students and lower-group students. In this case, it means that the number of UG students who answered correctly is not the same as the number of LG students who answered correctly.

Based on the findings of the data analysis, item difficulty and item discrimination of the test items in MAS Madani Pao-Pao Gowa were calculated into the easy category. In terms of item difficulty and discrimination, it caused several numbers including those in the simple category. The percentages of item difficulty and item discrimination of the multiple-choice test were 8 % in the difficult category, 28 % in the medium category, and 64% in easy category.

Based on the percentage of item discrimination on the multiple-choice test, of 39 students, there are 39% got the excellent category, 41% got the good category, 10% got the sufficient category, 5% got the insufficient category, and 5% got the poor category. Students who got insufficient and poor value, generally either did not understand English or had difficulties because of their cell phones. There was no test items in number that were excellent, good, and sufficient to discrimination the test 0%, insufficient needed to be reviewed (40%) or insufficient category while item numbers were poor and needed to be discarded (60%) or poor category. Besides that from 25 number items, there are 7 items, or 28% need to discard, while 18 items, or 72% need to review.

In other researchers, Hisbullah (2018) found that some test items were not excellent for discriminating the test, 3 test items were good and needed to be improved, there were 15 of the test items were sufficient, even though they still needed to be reviewed, there were 11 test items that were poor and need to be discarded, and 11 of the test items were very poor and needed correction right away. The percentage multiple-choice tests discriminate power, was tested in the good

category 7.5%, test in the satisfactory category 37.5 %, test in the poor category 27.5%, and test in the very poor category 27.5%. In contrast, Lestari (2011) found that out of 50 numbers, there were 35 numbers related to the category 0.25 to 0.75, or 70% of the items have positive discrimination criteria.

All research has advantages and disadvantages, and this researcher is no any different. This study only looked at the quality of students' responses to the English summative test at MAS Madani Pao-Pao Gowa, including the analysis of item difficulty and item discrimination for each item evaluated.

Based on the results of this research, item difficulty and item discrimination of the test items in MAS Madani Pao-Pao Gowa were counted into the easy category. The researcher recommended the teacher to perform a need analysis before providing instruction in the classroom and to conduct an analysis, review, and evaluation before giving a test to the students. In addition, the results of this research are only to find out whether the test items can discriminate between students, the upper group, and the lower group. Meanwhile, only two formulas were used, for the item difficulty and item discrimination formulas, because the questions were only multiple-choice.

CONCLUSION

The findings of this research that the test used in MAS Madani Pao-Pao Gowa was included into the good category. It that there 23 students got to exceed, there are 5 students got to achieve, and 11 students got not achieve or under minimum completeness criteria. There were two difficult items (9, and 14), or 8%, seven medium items (1, 2, 10, 11, 12, 20, 21), or 28%, and sixteen easy items (3, 4, 5, 6, 7, 8, 13, 15, 16, 17, 18, 19, 22, 23, 24, 25) or 64%. Based on the multiple-choice test's item discriminating percentage, of 39 students, there are 39% got the excellent category, 41% got the good category, 10% got the sufficient category, 5% got the insufficient category, and 5% got the poor category. Students who got insufficient and poor value, generally either did not understand English or had difficulties because of their cell phones. There were no test items that were excellent, good, or sufficient to differentiate the test 0%, insufficient needed to be reviewed (40%) or insufficient category while item numbers were poor and needed to be discarded (60%) or poor category. Besides that from 25 number items, there are 7 items, or 28% need to discard, while 18 items, or 72% need to review. Based on the results of this research, item difficulty and item discrimination of the test items in MAS Madani Pao-Pao Gowa were counted into the easy category. The researcher recommended the teacher to perform a need analysis before providing instruction in the classroom and to conduct an analysis, review, and evaluation before giving a test to the students.

REFERENCES

Baeti, A. N. (2020). *An Analysis of Student' Difficulties in Understanding Reading English Texts* [Bachelor, Universitas Muhammadiyah Purwokerto]. https://repository.ump.ac.id/12555/

Boopathiraj, C., & Chellamani, D. K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in The Test for Research In Education. *International Journal Of Social Sciences & Interdisciplinary Research*. https://www.semanticscholar.org/paper/ANALYSIS-OF-TEST-ITEMS-ON-DIFFICULTY-LEVEL-AND-IN-Boopathiraj-Chellamani/a9e08c03848e95760275e36f75cae88e49bc6c65

Brookhart, S. M., & Nitko, A. J. (2018). *Educational Assessment of Students* (Eighth). Perason.

Djiwadono, S. (2011). *Tes Bahasa: Pegangan bagi Pengajar Bahasa*. PT. Indeks.

Fadlilah, A. (2018). *An Analysis of Difficulty Level and Discriminating Power of English USBN Test 2018* [Undergraduate Thesis]. UIN Syarif Hidayatullah.

Hisbullah, H. (2018). *The Analysis of Discrimination Power of English Summative Test at MA Muhammadiyah Tengnga Lembang Sinjai* [Diploma, Universitas Islam Negeri Alauddin Makassar]. http://repositori.uin-alauddin.ac.id/12250/

Hotiu, A. (2006). *THE Relationship between Item Difficulty and Discrimination Indices in Multiple-Choice Tests in A Physical Science Course*. Florida Atlantic University.

Kocdar, S., Karadag, N., & Sahin, M. D. (2016). Analysis of the Difficulty and Discrimination Indices of Multiple-Choice Questions According to Cognitive Levels in an Open and Distance Learning Context. *TOJET: The Turkish Online Journal of Educational Technology*, *15*(4), 16–24.

Lestari, H. (2011). *An Item Analysis on Discriminating Power of English Summative Test (A Case Study of Second Year of SMPN 87 Pondok Pinang)*. https://repository.uinjkt.ac.id/dspace/handle/123456789/2826

Maharani, A. V., & Putro, N. H. P. S. (2020). Item Analysis of English Final Semester Test. *Indonesian Journal of EFL and Linguistics*, 491–504. https://doi.org/10.21462/ijefl.v5i2.302

Sengkaton, I., Hambali, M., & Mirizon, S. (2020). Teacher-made Summative Test: An Analysis of Test Format, Index of Difficulty, Discrimination, and Distractors. *LINGUA: Jurnal Bahasa Dan Sastra*, *20*(2), Article 2.