

Received: Mei 2024

Accepted: Juni 2024

Published: Juli 2024

Analysis of Different Power Levels, Difficulty, and Effectiveness of Distracting Final Semester Assessment Questions (PAS) for Arabic Language Class VIII MTs Irsyadul Athfal

Ahmad Zaini Dahlan, Fauzie Muhammad Shidiq, Zaky Anwar, Muhammad Syihabul Ihsan Al Haqiqy, Nur Qomari

Universitas Islam Negeri Maulana Malik Ibrahim Malang, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Universitas Islam Negeri Maulana Malik Ibrahim Malang

Email: zainidahlan1998@gmail.com, fauziemuhhammadshidiq@gmail.com, zakyanwarkarim@gmail.com, elhaqiqy123@gmail.com qomari@uin-malang.ac.id

Abstract

This study aims to determine the quality of the final exam items of Arabic language subjects in the 2023 academic year at MTs Irsyadul Athfal Gresik regarding the level of differentiation, difficulty, and effectiveness of exemptions. This research is an evaluation research with a quantitative descriptive approach. The research was conducted with a sample of 13 student answer sheets. The data collection technique used was the documentation technique. The results of this study indicate that (1) based on the criteria of differential power obtained 11 questions with inferior interpretation, 21 questions with poor interpretation, 1 with moderate interpretation, 9 with good interpretation, and 3 with very good interpretation. (2) based on the difficulty level criteria, 15% of the items were very difficult, 4% of the items were categorized as difficult, 2% of the items were moderate, 13% of the items were categorized as easy, and 64% of the items were very easy. (3) based on the effectiveness of the check obtained 1 item (2.22%) has a very good checker, then 34 items (75.55%) have a good enough checker or revised, and 10 items (22.22) that must be replaced because the checker does not work or rejected. The implication of this research is to inform the results of the analysis of the question items to be used as a comparison material for the test scores of Arabic language tests.

Keywords: Distinguishing power; Difficulty; Examination.

Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas butir soal ujian akhir semester mata pelajaran bahasa Arab pada tahun ajaran 2023 di MTs Irsyadul Athfal Gresik dilihat dari segi tingkat daya beda, kesukaran, dan keefektifan pengecoh. Penelitian ini merupakan penelitian yang bersifat

evaluasi dengan pendekatan deskriptif kuantitatif. Penelitian dilakukan dengan sampel sebanyak 13 lembar jawaban siswa. Teknik pengumpulan data yang digunakan adalah teknik dokumentasi. Hasil penelitian ini menunjukkan bahwa (1) berdasarkan kriteria daya beda diperoleh 11 soal dengan interpretasi jelek sekali, 21 soal dengan interpretasi jelek, 1 soal dengan interpretasi sedang, 9 soal dengan interpretasi baik, 3 soal dengan interpretasi baik sekali. (2) berdasarkan kriteria tingkat kesukaran diperoleh 15% butir soal sangat sukar, 4% butir soal dikategorikan sukar, 2% butir soal sedang, 13% butir soal yang dikategorikan mudah, dan 64% butir soal sangat mudah. (3) berdasarkan efektifitas pengecoh diperoleh 1 butir soal (2,22%) mempunyai pengecoh yang sangat baik, kemudian 34 butir soal (75,55%) mempunyai pengecoh yang cukup baik atau direvisi dan 10 butir soal (22,22) yang harus diganti karena pengecoh tidak berfungsi atau ditolak. Implikasi penelitian ini ialah untuk menginformasikan hasil analisis dari butir soal untuk bisa dijadikan bahan perbandingan nilai uji tes soal bahasa Arab.

Kata kunci: Daya beda, Kesukaran, Pengecoh.

Introduction

In general, tests have a very large role in teaching, including in Arabic language teaching. Even the test is in the most important position in it. In teaching theory, teaching can be seen as a process consisting of three main components that must be present in a teaching, first, teaching objectives second, teaching implementation third, assessment of teaching results.¹ Djiwandono mentions in his book that the assessment of teaching results is in the last position in the implementation of teaching and has a very urgent relationship and role with the previous components, one of which is the implementation of learning.² Because the results of the implementation of teaching can be known by using the assessment. Whether the goals in teaching are achieved or not achieved. So this last component is very essential in teaching, including Arabic language teaching.

Partially, the assessment is carried out to determine the achievement of learning objectives or to obtain an overview of the position of students in the flow of the learning process, namely regarding what students have mastered and what they still have to strive to master.³ In learning Arabic four skills must be mastered by learners (*listening skills, speaking skills, reading skills, and writing skills*). To find out the mastery of students in the four maharahs use assessment, referring to the competency standards in Arabic language learning that have been set.⁴ Therefore, it is necessary to assess the results of students at the end of the education unit. Assessment of the final results of students is carried out by the teacher through the end-of-semester exams.

The end-of-semester exam is an important evaluation that tests students' understanding and achievement in the

material that has been learned during a semester, in essence, the end-of-semester exam is an important moment that allows students to measure the extent to which they have understood the subject matter and how ready they are to move on to the next stage. Thus, the end-of-semester exam becomes one of the determining factors for students' academic achievement and a tool to evaluate teaching effectiveness.⁵

MTs Irsyadul Athfal is one of the educational units in Gresik that uses end-of-semester exams to determine the extent to which the teaching process and its achievements have been achieved. The end-of-semester assessment of Islamic Junior High School Irsyadul Athfal uses questions prepared by the Subject Teacher Consultation Team (MGMP), including Arabic learning material which is a clump in education based on Islamic religious education. Although the final semester exam questions are prepared by MGMP and labeled as national standards, the items need to be reviewed to find out how the level of differentiation, the level of difficulty, and the level of checking of the items that have been prepared in the Arabic language subject. If the items are not analyzed, it will not know the level of differentiation, difficulty, and checking of the questions tested to students. If no analysis of the items is carried out, then

the levels of the questions tested are not measurable and do not know their feasibility. Therefore, to find out the level of difficulty, differentiation, and exemption, it is necessary to study and review the question items quantitatively because the quantitative analysis of question items has to do with their statistical characteristics.⁶

According to Anastasi and Urbina, item analysis can be done qualitatively, about its content and form, as well as quantitatively, about its statistical characteristics.⁷ Qualitative analysis is more in terms of material, construction, language, and culture. While quantitative analysis is used to determine the characteristics of the items including the differentiation, difficulty, and exclusion of the items. Therefore, with the description of the problem above, the researcher is interested in researching the item analysis of the Arabic language map of class VIII MTs Irsyadul Athfal, which was prepared by MGMP with the national standard.

The question that arises is whether the Arabic questions in the End of Semester Assessment at MTs Irsyadul Athfal have met the evaluation criteria or can be used to measure a person's level of understanding. A participant is not considered a failure because they have not mastered the Arabic material taught.

Conversely, the level of differentiation, difficulty, and sophistication makes it difficult to understand the questions. In addition, if item analysis is not carried out, the quality of the items tested cannot be measured and the feasibility is unclear. Therefore, this study aims to answer the question of what is the quality of Arabic language subject questions used in the Final Semester Assessment Questions (PAS) for Arabic Language Class VIII MTs Irsyadul Athfa for the 2023/2024 academic year in terms of differentiability, difficulty level and effectiveness of checking.

The previous studies that support this research 1) Robi'atul Laili Maulidiyah et al. 2020 Development of Problem Items for Speaking Skills and Writing Skills in Class X students at the senior high school 1 Bojonegoro state Madrassa The results of his research were two designs of the development of the first question item about 25 items of taharah kalam and 25 questions of writing skills.⁸ Analysis of the Level of Difficulty and Distinguishing Power of Radiography Level 1 Training Exam Questions, the results are that there are 2 difficult items, 14 medium items, and 22 easy items. While the differentiating power is 7 general exam questions, 11 specific ones have low differentiating power. 3) Ratri Laksitaning Dewi. 2021. Analysis of the

level of difficulty of the final assessment questions in class V sports physical education lessons at Pangudang State Elementary School, Purwerejo district, the results are as follows, a total of 35 items, 18 items in the easy category, 13 items in the medium category, 4 items in the difficult category. The implication of this research is to inform the results of the analysis of the items to be used as a comparison material for the test scores of Arabic language tests.

Method

This research is an evaluation research with a quantitative descriptive approach with documentation data collection techniques.⁹ The data was taken from the Final Semester Assessment (PAS) items made by the Subject Teacher Conference (MGMP) which had been applied to class VIII students of MTs Irsyadul Athfal Gresik in the odd semester of the 2023/2024 school year, with a research sample consisting of 13 students. The data was analyzed from the aspects of differentiability, difficulty level, and checking with the data analysis technique of the test of differentiability, difficulty, and checking with the help of IBM SPSS Statistics 25 and Excel software.

Results and Discussion

Tests are an important part of the instructional decision-making process as they serve as tools that gather various information. Therefore, there will be many benefits from developing and using high-quality tests. A good exercise has features and characteristics that must be met, the characteristics of a quality test by applicable rules mainly include: (1) differentiating power, (2) level of difficulty, and (3) level of reliability.¹⁰ However, the characteristic beyond that is differentiating power, which is the property of a test item that shows that there is a difference between clever and less clever students. The higher the differentiating power of a test item, the more likely it is to distinguish clever students from less clever ones.

One of the next characteristics of a quality test is the level of difficulty of the test. The level of difficulty indicates how difficult or easy each part of the test is. Reviewing and analyzing the difficulty level of each part of the test can determine whether the test is too difficult, difficult, moderate, easy, or too easy.¹¹ In addition, the difficulty level of the exam can be calculated from the average score of the students. A high average score indicates that the exam was easy, while a low average score indicates that the exam was difficult or very difficult.

The next feature, examining the distribution of answer choices for the Final Semester Assessment Questions (PAS) for Arabic Language Class VIII MTs Irsyadul Athfal in 2023 is intended to determine whether the available answer choices are functional or not. An answer choice (exemplar) can be said to be functional if the exemplar: (1) at least 5% of test takers/students choose them, and (2) more are chosen by groups of students who do not understand the material.

Analysis of End of Semester Assessment Questions in terms of Differentiated Test Questions

The ability of a question to sort out high or superior learners from low-ability learners is called differentiating power.¹² The differentiating power of a question is the ability of a question item to distinguish students who have mastered the material being asked and students who have not mastered the material being asked. The differentiating power of a question, according to Aiken in Mutholib, is the ability of an item to distinguish between test takers who have learned and those who have not. The following are the benefits of item discriminating power. 1) To use empirical data to improve the quality of each part of the question. The discriminating power index can be used to determine whether each item is good, revised, or rejected. 2) To

determine how far each part of the question can identify or differentiate student abilities, including students who have understood or not understood the material taught by the teacher. If the question part cannot distinguish between the two abilities of students, then the question part can be suspected for several reasons, such as the answer key for the question part is incorrect; the answer key for the question part has two or more correct answer keys.

A question with a higher discriminating power index indicates that the

question is better able to distinguish students who have understood the material from students who have not understood it. The discriminating power index ranges from -1.00 to +1.00. The higher the differentiating power of a question, the stronger/better it is. If the discriminating power is negative (<0), it means that there are lower groups (students who do not understand the material). Based on the above starting point, there is a standard that can be used to determine how much of a difference each particular question item can make as the following standards.¹³

Table 1: Discrimination Index

The magnitude of the item discrimination index number (D)	classification	Interpretation
Less than 0.20	Poor	The item in question has very weak differentiating power (bad) and is considered not to have good differentiating power.
0,20 - 0,40	Satisfactory	The item in question has sufficient differentiating power (medium).
0,40 - 0,70	Good	The item in question has good discriminating power.
0,70 - 1,00	Excellent	The item in question has excellent discriminating power.
Negative Sign	-	The item in question has negative discriminating power (very bad).

To find out the differentiating power of multiple-choice questions, use the following formula.

$$DP = \frac{BA - BB}{\frac{1}{2}N} \text{ Atau } \frac{2(BA - BB)}{N}$$

DP = question differentiating power,

BA = number of correct answers in the upper group

BB = number of correct answers in the lower group,

N = number of students who took the test taking the test.

The differentiating power obtained of the 2023/2024 academic year using SPSS from the item analysis of Arabic multiple choice questions in class VIII odd semester 26 is as follows.

Table 2: Distinguishing Power

Item-Total Statistics					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Interpretation
1	31,7	8,4	0,00	0,49	Bad
2	32,0	8,5	-0,12	0,52	Very Bad
3	31,8	8,0	0,19	0,48	Bad
4	31,8	8,0	0,19	0,48	Bad
5	31,8	7,4	0,64	0,43	Good
6	31,8	6,6	0,83	0,37	Very Good
7	31,8	6,6	0,83	0,37	Very Good
8	31,8	7,7	0,41	0,45	Good
9	32,7	8,4	0,00	0,49	Bad
10	31,7	8,4	0,00	0,49	Bad
11	31,8	6,6	0,83	0,37	Very Good
12	31,8	7,4	0,64	0,43	Baik
13	32,1	8,9	-0,25	0,55	Very Bad
14	31,8	8,0	0,19	0,48	Bad
15	31,8	7,8	0,21	0,47	Bad
16	31,7	8,4	0,00	0,49	Bad
17	31,8	7,6	0,30	0,46	Medium
18	31,7	8,4	0,00	0,49	Bad
19	31,7	8,4	0,00	0,49	Bad
20	31,8	7,7	0,41	0,45	Good
21	31,7	8,4	0,00	0,49	Bad
22	32,3	8,4	-0,09	0,52	Very Bad
23	32,4	8,4	-0,09	0,52	Very Bad
24	31,7	8,4	0,00	0,49	Bad
25	31,7	8,4	0,00	0,49	Bad
26	31,7	8,4	0,00	0,49	Bad
27	32,0	8,8	-0,23	0,54	Very Bad
28	32,0	7,8	0,12	0,48	Bad
29	32,5	10,1	-0,77	0,59	Very Bad
30	32,7	8,4	0,00	0,49	Bad
31	31,7	8,4	0,00	0,49	Bad
32	32,6	8,9	-0,36	0,53	Very Bad
33	31,8	7,4	0,64	0,43	Baik
34	31,8	8,0	0,19	0,48	Bad

35	32,6	8,9	-0,36	0,53	Very Bad
36	32,7	8,4	0,00	0,49	Bad
37	31,8	7,3	0,47	0,43	Good
38	31,8	7,4	0,64	0,43	Good
39	32,6	8,9	-0,36	0,53	Very Bad
40	31,7	8,4	0,00	0,49	Bad
41	31,8	7,3	0,47	0,43	Good
42	31,8	7,0	0,65	0,40	Good
43	31,8	8,1	0,05	0,49	Bad
44	32,2	9,2	-0,34	0,57	Very Bad
45	32,1	8,9	-0,25	0,55	Very Bad

The differentiating power of the items is contained in the Corrected Item-Total Correlation table. In the Arabic multiple choice questions of class VIII odd semester of the 2023/2024 academic year, 11 questions were found with the interpretation of "very bad", namely the question is not suitable for use because the item concerned has negative differentiating power, 21 questions with the interpretation of "bad" are considered not to have good differentiating power, 1 question with a "medium" interpretation, which means that the item concerned has sufficient differentiating power, 9 questions with a "good" interpretation, namely the item concerned has good differentiating power, and finally 3 questions with a "very good" interpretation, which means that the item concerned has very good differentiating power.

Analysis of End of Semester Assessment Questions in terms of Level Difficulty

According to Asmawi Zainul, in ¹⁴

The percentage of examinees who answer a particular question correctly is called item difficulty. The level of item difficulty is usually represented by a p-value, where a larger p-value indicates that the level of item difficulty is lower. This means the problem is getting easier, and vice versa. To determine the level of difficulty of questions in Arabic language subjects, we must consider the value of the proportion of examinees who answer each item correctly. In the Year 2023 End of Semester Assessment, the following steps were taken:

- 1) Observing and correcting the results of the test takers' answers.
- 2) Entering each test taker's answer item into the data view of the SPSS application, 1 for the correct category and 0 for the wrong category, and calculating the scores for each item.

3) Then the calculation is done by finding the average/mean. In general, using the formula= $P \frac{B}{JS}$. Where P = difficulty index, B = number of students answering correctly, JS = total number of students.

4) Furthermore, from the output of the calculation of SPSS Version 25 on the Level of Difficulty of the Arabic language items of the End of Semester Assessment of the odd semester of 2023, the numbers are obtained as stated in the following table.

Table 3. SPSS Output Results Of Item Analysis Of Arabic Language Questionnaire Final Assessment Odd Semester 2023

		soal1	soal2	soal3	soal4	soal5	soal6	soal7	soal8	soal9	soal10
N	Valid	13	13	13	13	13	13	13	13	13	13
	Missing	0	0	0	0	0	0	0	0	0	0
Mean		1,0000	0,6923	0,9231	0,9231	0,9231	0,8462	0,8462	0,9231	0,0000	1,0000
soal11	soal12	soal13	soal14	soal15	soal16	soal17	soal18	soal19	soal20	soal21	soal22
13	13	13	13	13	13	13	13	13	13	13	13
0	0	0	0	0	0	0	0	0	0	0	0
0,7692	0,9231	0,6923	0,9231	0,9231	1,0000	0,8462	0,8462	1,0000	0,9231	1,0000	0,3846
soal23	soal24	soal25	soal26	soal27	soal28	soal29	soal30	soal31	soal32	soal33	soal34
13	13	13	13	13	13	13	13	13	13	13	13
0	0	0	0	0	0	0	0	0	0	0	0
0,3077	1,0000	1,0000	1,0000	0,6923	0,6923	0,1538	0,0000	1,0000	0,0769	0,9231	0,9231
soal35	soal36	soal37	soal38	soala39	soal40	soal41	soal42	soal43	soal44	soal45	
13	13	13	13	13	13	13	13	13	13	13	
0	0	0	0	0	0	0	0	0	0	0	
0,0769	0,0000	0,8462	0,9231	0,0769	1,0000	0,8462	0,8462	0,8462	0,4615	0,6154	

From the results of the data output view SPSS can be seen in the mean table, the results of the mean were consulted with the percentage (item difficulty index / IKB). Worth 0.00- 1.00; with IKB criteria: 0.00-

0.20 is very difficult; 0.20-0.40 difficult; 0.40-0.60 medium; 0.60-0.80 easy, and 0.80-1.00 very easy.¹⁵ As can be classified by researchers in the form of tabulation as follows:

Table 4. Consultation Results of Difficulty Index

No.	Results	Difficulty Level (%)	Description
1	100	100	very easy
2	69	69	easy
3	92	92	very easy
4	92	92	very easy
5	92	92	very easy
6	84	84	very easy

7	84	84	very easy
8	92	92	very easy
9	0	0	very difficult
10	100	100	very easy
11	76	76	easy
12	92	92	very easy
13	69	69	easy
14	92	92	very easy

15	92	92	very easy	31	100	100	very easy
16	100	100	very easy	32	0	0	very difficult
17	84	84	very easy	33	92	92	very easy
18	84	84	very easy	34	92	92	very easy
19	100	100	very easy	35	0	0	very difficult
20	92	92	very easy	36	0	0	very difficult
21	100	100	very easy	37	84	84	very easy
22	38	38	difficult	38	92	92	very easy
23	30	30	difficult	39	0	0	very difficult
24	100	100	very easy	40	100	100	very easy
25	100	100	very easy	41	84	84	very easy
26	100	100	very easy	42	84	84	very easy
27	69	69	Easy	43	84	84	very easy
28	69	69	easy	44	46	46	medium
29	15	15	very difficult	45	61	61	easy
30	0	0	very difficult	Average		71	

From this table it can be concluded that there are 7 items of Arabic subject matter for the End of Odd Semester Assessment in 2023 that have a Level of Difficulty (LD) between 0.00-0.20: very difficult category and as many as 2 questions have LD between 0.20-0.40: difficult

category. 1 question has TK between 0.40-0.60: medium category. 6 items have LD between 0.60-0.80: easy category, and as many as 29 questions have LD between 0.80-1.00: very easy category. The details of the status of these items are as follows:

Table 5. Status Of Question Items

Question Status	Question Item Number	Total
very difficult	9, 29, 30, 32, 35, 36, 39	7
difficult	22, 23	2
medium	44	1
easy	2, 11, 13, 27, 28, 45	6
Very easy	1, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 16, 17, 18, 19, 20, 21, 24, 25, 26, 31, 33, 34, 37, 38, 40, 41, 42, 43	29

This means that if it is projected, there are 15% of the items from the final assessment of the odd semester of Arabic language subjects in 2023 categorized as very difficult; 4% of the items are categorized as difficult; 2% of the items are

Medium; 13% of the items are categorized as easy, and 64% of the items are very easy.

And when viewed the average value of the level of difficulty for all items of the final assessment of the odd semester of Arabic subject 2023 is 71%. Thus it can be

said that the Arabic language subject matter of the odd PAS in 2023 has an easy level of difficulty and cannot be used as a standardized test. A good question is not too easy or too difficult, according to Suharsimi Arikunto. Questions that are too easy will not make students try harder to solve them, while questions that are too difficult will make students desperate and not eager to try again because they are out of reach.¹⁶.

Analysis Of Odd Semester Final Assessment Questions In Terms Of Testability (Distractor)

One way to find out how effective an exemplar is by looking at the pattern of student answer distribution. The answer distribution pattern can be calculated by counting how many test takers choose one of the options or do not choose at all. If this happens, an exception is considered effective if at least 5% of all test takers have chosen that option. The more test takers who select an exception, the better the exception works.¹⁷. Learners omit if they ignore all options (do not select). The test is said to be good if the omission is no more than 10% of the learners. Information on reading the level of checking power on the question answer items as follows:

An option is said to be good if it has the power of distraction (exception)

1. At least 5% of the learners are selected.
2. The group of less intelligent students chooses more than the group of intelligent students.

Formula: $D = A/N \times 100\%$

Description:

D: Distraction Level (%)

A: Number of Students who choose that Option

N: Total number of students

The criteria are as follows:

- If $D \geq 5\%$ Then accepted because it is good
- If $D \leq 5\%$ and ≤ 0 Then it is revised/rewritten because it is less good
- If $D = 0$ then rejected because it is not good

Table 6. Distractor level

Distractor Level										
ITEMS	SOAL 1	SOAL 2	SOAL 3	SOAL 4	SOAL 5	SOAL 6	SOAL 7	SOAL 8	SOAL 9	SOAL 10
A	0%	15%	8%	0%	0%	0%	0%	8%	0%	0%
B	0%	8%	92%	92%	92%	0%	85%	92%	0%	100%
C	0%	69%	0%	8%	8%	85%	8%	0%	85%	0%
D	100%	8%	0%	0%	0%	15%	8%	0%	15%	0%
Distractor Level										
ITEMS	SOAL 11	SOAL 12	SOAL 13	SOAL 14	SOAL 15	SOAL 16	SOAL 17	SOAL 18	SOAL 19	SOAL 20
A	85%	8%	0%	92%	0%	0%	0%	0%	100%	92%
B	8%	0%	62%	0%	92%	100%	0%	0%	0%	0%
C	8%	8%	8%	8%	0%	0%	15%	92%	0%	8%
D	0%	85%	31%	0%	8%	0%	85%	8%	0%	0%
Distractor Level										
ITEMS	SOAL 21	SOAL 22	SOAL 23	SOAL 24	SOAL 25	SOAL 26	SOAL 27	SOAL 28	SOAL 29	SOAL 30
A	0%	0%	0%	100%	100%	100%	0%	23%	0%	15%
B	0%	38%	69%	0%	0%	0%	69%	69%	77%	0%
C	0%	62%	0%	0%	0%	0%	31%	8%	0%	8%
D	100%	0%	31%	0%	0%	0%	0%	0%	23%	77%
Distractor Level										
ITEMS	SOAL 31	SOAL 32	SOAL 33	SOAL 34	SOAL 35	SOAL 36	SOAL 37	SOAL 38	SOAL 39	SOAL 40
A	0%	23%	0%	85%	15%	92%	85%	0%	8%	0%
B	0%	69%	92%	8%	0%	0%	15%	8%	0%	0%
C	0%	8%	8%	0%	85%	8%	0%	0%	92%	100%
D	100%	0%	0%	8%	0%	0%	0%	92%	0%	0%
ITEMS	SOAL 41	SOAL 42	SOAL 43	SOAL 44	SOAL 45					
A	8%	85%	85%	54%	69%					
B	85%	0%	0%	46%	31%					
C	0%	8%	15%	0%	0%					
D	8%	8%	0%	0%	0%					

There should be a difference in the frequency of answers between upper-group students and lower-group students.¹⁸ It is known Based on the analysis of the effectiveness of the triggers that have been carried out, most of the triggers on the Final Examination of the Odd Semester of

MTs Irsyadul Athfal Jatirembe Benjeng Gresik are not functioning properly because there is only 1 item (2.22%) that has a very good trigger, then 34 items (75.55%) have good enough triggers or are revised and 10 items (22.22) that must be replaced because the triggers do not work or are rejected.¹⁹

An exception is said to be functional if it is chosen by at least 5% of the number of students who take the test and is chosen by students who do not master the test material.²⁰ Can be said to be misleading if the exception is chosen by most students²¹. Of the 45 items tested, there were 4 questions with misleading exceptions, namely numbers 9, 35, 36, and 39.

Conclusion

The results of the analysis of the differential power on the multiple-choice End-of-Semester Assessment questions of Arabic language subjects in class VIII odd semesters of MTs Irsyadul Athfal in the 2023/2024 academic year, found 11 questions with the interpretation of "very bad", namely the question is very unfit for use because the item concerned has negative differentiating power, 21 questions with a "bad" interpretation are considered not to have good differentiating power, 1 question with a "medium" interpretation, which means that the item concerned has sufficient differentiating power, 9 questions with a "good" interpretation, namely the item concerned has good differentiating power, and finally 3 questions with a "very good" interpretation, which means that the item concerned has very good differentiating power.

As for the level of difficulty, if it is projected, there are 15% of Arabic language

subject items for the Odd Semester Final Assessment in 2023, categorized as very difficult; 4% of the items are categorized as difficult; 2% of the items are medium; 13% of the items are categorized as easy, and 64% of the items are very easy. When viewed, the average value of the level of difficulty for all Arabic language subject matter items for the Odd Semester Final Assessment in 2023 is 71%. Thus it can be said that the Arabic language subject matter of the Odd Semester Final Assessment in 2023 has an easy level of difficulty, in the sense that it cannot be used as a standard test. For the level of checking is not functioning properly because there is only 1 item (2.22%) that has a very good checker, then 34 items (75.55%) have a fairly good checker or are revised, and 10 items (22.22) that must be replaced because the checker does not work or is rejected.

References

- Adawiyah, Indzar Robiatul, and Syarifuddin. "Pengaruh Media Kahoot Terhadap Peningkatan Pembelajaran Maharah Istima' Di MTs Wali Songo Purwosari Pasuruan." *Al-Kalim: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban* 2, no. 2 (2023): 135–49.
- Agustiawan, Eko, Anies Fuady, and Sunismi. "Analisis Butir Soal Tes Uraian Ujian Tengah Semester Mata Kuliah Statistik." *EDUCARE: Jurnal Ilmu*

- Pendidikan* 02, no. 01 (2023): 16–24.
- Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Edisi 2. Jakarta: PT Bumi Aksara, 2014.
- Aryobimo, Bagusradityo. “Analisis Butir Soal Latihan Pada Buku Al-Naḥw Al-Wāḍiḥ Untuk Siswa Madrasah Aliyah.” *Al-Ma’rifah* 18, no. 2 (2021): 127–38.
<https://doi.org/10.21009/almakrifah.18.02.02>.
- Azis, Ahmad Miftachul. *Analisis Butir Soal UAS Gasal 2017/2018 Muatan Bahasa Indonesia Kelas IV SD*. JURNAL PENDIDIKAN: Riset dan Konseptual, n.d.
- Bramantha, Heldie, and Siti Rahmania. “Analisis Butir Soal Berdasarkan Tingkat Kesukaran Dan Daya Pembeda Pada Buku Siswa Tematik Terpadu Sekolah Dasar” 10, no. 2 (2022): 451–57.
<https://doi.org/10.25273/jems.v10i2.15130>.
- Febri, Retno Inten Fitria, Achmad Arwan, and Faizatul Amalia. “Pengembangan Sistem Analisis Butir Soal Ujian Berbasis Web (Studi Kasus: SMP Negeri 1 Donomulyo).” *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer* 5, no. 3 (2021): 943–49.
- Halik, Andi Surahma, Sitti Mania, and Fitriani Nur. “Analisis Butir Soal Ujian Akhir Sekolah (Uas) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 Smp Negeri 36 Makassar.” *Al Asma : Journal of Islamic Education* 1, no. 1 (2019): 11.
<https://doi.org/10.24252/asma.v1i1.11249>.
- Hasbullah, Habib. “Analisis Butir Soal Matematika UN SMP/MTs Tahun 2018/2019 Ditinjau Dari Kriteria Koneksi Matematis.” *Efektor* 7, no. 1 (2020): 73–79.
<https://doi.org/10.29407/e.v7i1.1436>.
- Khofifah, Utami, Jujun Ratnasari, and Aa Juhanda. “Analisis Butir Soal Kemampuan Tree Thinking Pada Instrumen Uji Coba Materi Spermathophyta.” *Biodik* 6, no. 3 (2020): 328–35.
<https://doi.org/10.22437/bio.v6i3.9907>.
- Laela Umi Fatimah, Khairuddin Alfath. “Analisis Kesukaran Soal Daya Pembeda Dan Fungsi Distraktor” 8 (2019): 37–64.
- Lapenia, Marisa, and Hidayati. “Analisis Butir Soal Ujicoba Posttest Untuk Mengukur Hasil Belajar Dan Pemahaman Siswa Pada Materi Fluida.” *Jurnal Pendidikan Tambusai* 7, no. 3

(2023): 22935–43.

Lestari, Agistia Suci, Aflich Yusnita Fitrianna, and Luvy Sylviana Zanthly. “Analisis Butir Soal Tes Materi Sistem Persamaan Linear Dua Variabel Pada Siswa Kelas VIII.” *Jurnal Pembelajaran Matematika Inovatif* 6, no. 1 (2023): 367–76.

<https://doi.org/10.22460/jpmi.v6i1.12389>.

Mochammad Noor Akhmadi. “Analisis Butir Soal Evaluasi Tema 1 Kelas 4 Sdn Plumbungan Menggunakan Program Anates.” *Ed-Humanistics : Jurnal Ilmu Pendidikan* 6, no. 1 (2021): 799–806. <https://doi.org/10.33752/ed-humanistics.v6i1.1464>.

Nur Cahyo, Ahmad, Deby Luriawati, and Wagiran Wagiran. “Analisis Butir Soal Penilaian Keterampilan Kebahasaan Pada Pembelajaran Teks Eksplanasi Kelas XI.” *JBSI: Jurnal Bahasa Dan Sastra Indonesia* 2, no. 01 (2022): 11–22. <https://doi.org/10.47709/jbsi.v2i01.1493>.

Okryanida, Indica Yona, Sri Mayanty, and Fita Widiyatun. “Analisis Butir Soal Kemampuan Berpikir Kritis Siswa SMAIT Nururrohman Depok.” *Jurnal Penelitian Pembelajaran Fisika* 15, no. 1 (2024): 73–79. <https://doi.org/10.26877/jp2f.v15i1.1>

7057.

Purniasari, Laksmi, Mohammad Masykuri, and Sri Retno Dwi Ariani. “Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019/2020 Menggunakan Model Iteman Dan Rasch.” *Jurnal Pendidikan Kimia* 10, no. 2 (2021): 205–14.

Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif Dan Kombinasi (Mixed Methods)*. Bandung: Alfabeta, 2020.

Virginia, Sagita, Widiya Angraini, Wika Pratesya, and Ahmad Walid. “Analisis Butir Soal Ulangan Harian IPA Terpadu Kelas VII SMP 05 Kota Bengkulu Tahun Pelajaran 2020/2021.” *Jurnal PTK Dan Pendidikan* 6, no. 2 (2021): 61–65. <https://doi.org/10.18592/ptk.v6i2.4173>.

Wirandani, Tari, Ayu Cendra Kasih, and Latifah. “Analisis Butir Soal Hots (High Order Thinking Skill) Pada Soal Ujian Sekolah Kelas XII Mata Pelajaran Bahasa Indonesia Di Smk an-Nahl.” *Parole:Pendidikan Bahasa Dan Sastra Indonesia* 2, no. 4 (2019): 488.

Endnote

¹ Wirandani, Kasih, and Latifah, "Analisis Butir Soal Hots (High Order Thinking Skill) Pada Soal Ujian Sekolah Kelas XII Mata Pelajaran Bahasa Indonesia Di Smk an-Nahl."

² Lopenia and Hidayati, "Analisis Butir Soal Ujicoba Posttest Untuk Mengukur Hasil Belajar Dan Pemahaman Siswa Pada Materi Fluida."

³ Hasbullah, "Analisis Butir Soal Matematika UN SMP/MTs Tahun 2018/2019 Ditinjau Dari Kriteria Koneksi Matematis."

⁴ Virginia et al., "Analisis Butir Soal Ulangan Harian IPA Terpadu Kelas VII SMP 05 Kota Bengkulu Tahun Pelajaran 2020/2021."

⁵ Nur Cahyo, Luriawati, and Wagiran, "Analisis Butir Soal Penilaian Keterampilan Kebahasaan Pada Pembelajaran Teks Eksplanasi Kelas XI."

⁶ Khoififah, Ratnasari, and Juhanda, "Analisis Butir Soal Kemampuan Tree Thinking Pada Instrumen Uji Coba Materi Spermathophyta."

⁷ Halik, Mania, and Nur, "Analisis Butir Soal Ujian Akhir Sekolah (Uas) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 Smp Negeri 36 Makassar."

⁸ Adawiyah and Syarifuddin, "Pengaruh Media Kahoot Terhadap Peningkatan Pembelajaran Maharah Istima' Di MTs Wali Songo Purwosari Pasuruan."

⁹ Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif Dan Kombinasi (Mixed Methods)*.

¹⁰ Aryobimo, "Analisis Butir Soal Latihan Pada Buku Al-Nahw Al-Wāḍiḥ Untuk Siswa Madrasah Aliyah."

¹¹ Lestari, Fitrianna, and Zanthi, "Analisis Butir Soal Tes Materi Sistem Persamaan Linear Dua Variabel Pada Siswa Kelas VIII."

¹² Agustiawan, Fuady, and Sunismi, "Analisis Butir Soal Tes Uraian Ujian Tengah Semester Mata Kuliah Statistik."

¹³ Laela Umi Fatimah, "Analisis Kesukaran Soal Daya Pembeda Dan Fungsi Distraktor."

¹⁴ Purniasari, Masykuri, and Ariani, "Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019/2020 Menggunakan Model Iteman Dan Rasch."

¹⁵ Okyranida, Mayanty, and Widiyatun, "Analisis Butir Soal Kemampuan Berpikir Kritis Siswa SMAIT Nururrohmah Depok."

¹⁶ Arikunto, *Dasar-Dasar Evaluasi Pendidikan*.

¹⁷ Bramantha and Rahmania, "Analisis Butir Soal Berdasarkan Tingkat Kesukaran Dan Daya Pembeda Pada Buku Siswa Tematik Terpadu Sekolah Dasar."

¹⁸ Febri, Arwan, and Amalia, "Pengembangan Sistem Analisis Butir Soal Ujian Berbasis Web (Studi Kasus: SMP Negeri 1 Donomulyo)."

²⁰ Azis, *Analisis Butir Soal UAS Gasal 2017/2018 Muatan Bahasa Indonesia Kelas IV SD*.

²¹ Mochammad Noor Akhmadi, "Analisis Butir Soal Evaluasi Tema 1 Kelas 4 Sdn Plumbungan Menggunakan Program Anates."