

An Analysis of Final Semester Exam Questions for the Arabic Language Subject of Merdeka Curriculum at SMA 4 Muhammadiyah Jakarta

M Taufiq Arkan¹, Ahmad Rizki Nugrahwan¹

¹ *Universitas Muhammadiyah Prof. Dr. Hamka (Uhamka) Jakarta, Indonesia*

 Taufiqarkan11@gmail.com

Abstract

This research aims to determine the quality of odd semester final semester questions for Arabic language subjects in the Independent Curriculum at SMA 4 Muhammadiyah Jakarta, which includes validity, reliability, level of difficulty, differentiability, distractibility, Pap, Pan, and suitability of the questions to the demands of the Independent Curriculum. Method used in this research is mixed methods, with quantitative and qualitative descriptive research characteristics. This research is a descriptive study that describes and interprets 40 multiple choice questions. Data was obtained using the documentation method, namely by copying the questions, answer keys and student answer sheets and then analyzed. The research results showed that: almost all the questions were in the valid category, and only a few questions were in the invalid criteria. in terms of reliability, it shows high reliability. In terms of level of difficulty, it shows that this question has 3 indices, namely difficult, easy and medium. In terms of differentiating power, it shows that you have questions with a good index. in terms of distracting power, it shows that more options are accepted compared to revisions. Student scores have been converted to PAP and show that the majority of children got A and B grades while only a small percentage got C and D grades. Based on the conversion of PAN scores, it can be shown that the majority of children get grades B and C, for grades A, D and E only some children. Thus, overall the final exam questions for class X in the field of Arabic language studies are sufficient and worthy of being given to students. In future, questions that are not good or distractors that don't work need to be corrected so that future exams will be better.

Keywords: Arabic Language, Merdeka Curriculum, Analysis of Final Semester Exam

ARTICLE INFO

Article history:

Received

January 03, 2024

Revised

March 27, 2024

Accepted

April 05, 2024

Published by
ISSN

Website

This is an open access article under the CC BY SA license

CV. Creative Tugu Pena
2774-4299

<https://attractivejournal.com/index.php/bse/>

<https://creativecommons.org/licenses/by-sa/4.0/>



INTRODUCTION

Education involves the entire series of life experiences and all forms of human interaction with the environment with the aim of optimizing the development of individual potential according to stages of development (Putri et al., 2023). The main goal is to reach a specific level of maturity (Tarmizi et al., 2021). In the context of personality development, education is considered a basic need (Pardimin et al., 2017). Education has a significant role in forming a person's character, and is recognized as a force that will have a direct impact on individual achievement and productivity (Sari et al., 2022). Through the

help of education, a person can understand and interpret the surrounding environment, enabling him to create extraordinary works in his life's journey (Genetics & Khairurrijal, n.d.). In education there are many curricula, one of which is the independent curriculum.

The Independent Curriculum is a learning method that has variations in intracurricular learning. This approach places special emphasis on the material so that students have sufficient time to learn concepts and strengthen their abilities. Teachers have the freedom to choose a variety of educational tools, allowing learning to be tailored to students' needs and desires. Projects aimed at raising the profile of Pancasila students are developed based on certain themes set by the government. They do not focus on specific learning objectives, so they are not tied to certain subject matter (Marsela Yulianti et al., 2022). In the independent curriculum there is an evaluation aspect.

Student development can be identified through evaluation of the teaching and learning process (Tilaar & Hasriyanti, 2019). Assessment is a general evaluation tool used to assess how well students understand the lessons they have been taught (Ida & Musyarofah, 2021). In addition, through analysis of the tests used, it will be assessed to what extent the tests reflect the achievement of curriculum objectives. Thus, the evaluation results can be used as an indicator of success in achieving learning objectives (Hamimi et al., 2020). One aspect of evaluation is educational assessment.

Educational assessment is an inseparable process from learning activities which aims to evaluate the achievement of educational goals (Fitrianawati, 2015). This evaluation is carried out with the aim of assessing the extent to which educational goals can be achieved (Nurjanah & Marliansih, 2015). The purpose of evaluation is to measure the extent to which students understand the lesson so that they can measure their level of intelligence (Fathiyah, 2019). Furthermore, evaluation is also useful for teachers and educational institutions to assess the success of the learning process and evaluate the effectiveness of the overall educational process in schools (Hendrayani, 2016). Success in the teaching and learning process can be assessed from the application of teaching methods, and one aspect that attracts attention is the choice of learning media used by educators in the learning context (Rachmawati, 2021). And exams are one aspect of an assessment.

A test is considered effective as a measuring tool if it meets the criteria of reliability, validity, differentiability of questions, level of difficulty, pap and pan (Hidayati & Nisa', 2023). In addition, exam questions must be in accordance with the applicable curriculum content (Halik et al., 2019). because the curriculum will be a reference for all teachers for the teaching and learning process, ensuring that the material taught is in accordance with the implemented curriculum (Alfarisa & Purnama, 2019). If the exam does not reflect the basic competencies listed in the curriculum, then it will not be effective as an evaluation tool (Halik et al., 2019). For this reason, research must be carried out on final semester exam questions to determine whether these questions meet the criteria for being effective questions, because the quality of questions can have an impact on student learning outcomes (Hendrayani, 2016). And one way to find out if a test is effective is to analyze the question items.

One of the actions that teachers must take to improve the quality of the questions they create is to carry out detailed analysis of the questions. Often, assessors forget to carry out evaluation tasks on the measuring tools used to assess student learning success (Mania et al., 2020). Exam quality analysis is a process that must be carried out to assess the quality of the exam (Kurniawan et al., 2017). According to (Muluki, 2020) "Item analysis or item analysis is a review of test questions to ensure that the set of questions is of adequate quality." "Question analysis is a systematic procedure that provides very specific information regarding the test items that we prepare," said (Jufni & Ibrahim, 2015). One method used to process test results to improve the quality of the learning process is item analysis (Nayla Amalia & Widayati, 2012).

Based on the views of the experts above, the author concludes that a teacher must carry out item analysis to assess the quality of each question they give. Through this analysis activity, teachers can gain a better understanding of which tests are good and worth keeping and which ones should be eliminated.

Previous researchers have conducted a lot of research on the difficulty of analyzing test items. The first research conducted by (Fikriyah, 2021) was entitled Analysis of Mid-Semester Exam Questions for Arabic Language Subject Class VII Odd Semester Muhammadiyah Middle School 1 Yogyakarta 2019/2020 Academic Year. This research aims to determine the suitability and quality of Mid-Term Exam question items for Subjects. Arabic Language Class VII Even Semester SMP Muhammadiyah 1 Yogyakarta Academic Year 2019/2020. The second research was conducted by (Maulana & Sanusi, 2020) with the title Analysis of Arabic Language Question Items for the Joint Regional Madrasah Final Examination (UAMBD) Madrasah Ibtidaiyah Year 2017-2018 Research This aims to analyze the Arabic language question items in the joint regional madrasa final exam (UAMBD) for Madrasah Ibtidaiyah (MI) students for the 2017-2018 academic year. The third research was conducted by (Makhisoh & Suroyya, 2022) with the title Analysis of Arabic Madrasah Exam Question Items According to Bloom's Taxonomy, this research focuses on the suitability of each Arabic final exam question item with the Madrasah Exam grid set by the Ministry of Religion in 2021 as well as measuring the level of validity of the questions as a benchmark for the suitability of the assessment instrument combined with Bloom's Taxonomy.

From the three examples of research above, we find similarities about this research. Both of them conducted research related to the analysis of question items which were evaluated in terms of validity, rehabilitation, level of difficulty and differentiating power. However, the novelty or difference with this research is in terms of location and the questions to be analyzed. We will test whether the quality of the SMA 4 Muhamadiya Jakarta year tests is appropriate and appropriate for students, and also for test development for the coming year. That is the novelty of the research that we conducted with questions and from different schools in order to create efficient learning evaluations and quality questions.

METHOD

This research adopts a mixed methods approach, with quantitative and qualitative descriptive research characteristics. This research adopts a mixed methods approach, with quantitative and qualitative descriptive research characteristics. There are several techniques that can be used to analyze question items qualitatively, including moderator techniques and panel techniques (Utomo, 2019). Quantitative question review is the study of question items based on empirical data. This empirical data is obtained from questions that have been tested. There are two approaches to quantitative analysis, namely the classical and modern approaches.

RESULTS AND DISCUSSION

1. Level of difficulty

A quality test is a test that avoids difficulty levels that are too low or high (Sappaile, Baso, 2007). Apart from considering the validity and reliability of the test, it is also necessary to check the difficulty level of the test to assess its quality and ensure that the difficulty level corresponds to the difficulty coefficient of each question item. (Statistics, 2023)

QUESTION ITEM NO	INDEX OF DIFFICULTY	CATEGORY
1	0,78	Difficult
2	0,58	Currently
3	0,74	Difficult
4	0,85	Difficult
5	0,22	Easy
6	0,62	Currently
7	0,49	Currently
8	0,60	Currently
9	0,64	Currently
10	0,64	Currently
11	0,52	Currently
12	0,43	Currently
13	0,38	Currently
14	0,49	Currently
15	0,29	Easy
16	0,28	Easy
17	0,47	Currently
18	0,37	Currently
19	0,52	Currently
20	0,48	Currently
21	0,54	Currently
22	0,52	Currently
23	0,44	Currently
24	0,48	Currently
25	0,53	Currently
26	0,40	Currently
27	0,48	Currently
28	0,74	Difficult
29	0,31	Currently
30	0,37	Currently
31	0,51	Currently
32	0,54	Currently
33	0,57	Currently
34	0,76	Difficult
35	0,61	Currently
36	0,55	Currently
37	0,54	Currently
38	0,49	Currently
39	0,75	Difficult
40	0,62	Currently

The following are the results of calculating the level of difficulty of the questions for each item in the multiple choice question type. The calculation results and criteria for the level of difficulty of the questions in each item of the multiple choice question type are presented in the table. The level of difficulty shows that 5(12.5%) questions have a difficult index, 3(7.5%) questions have an easy index, and 32(80%) questions have a medium index.

2. Differentiating power

Discriminating power is a student's ability to distinguish students who are very proficient or master the material from students who do not (Saputra et al., 2022). clever student. According to (Verawati et al., 2023) question analysis becomes relevant when the focus is on analyzing differentiating power. Differentiating power is used to separate students who have been able to master the lesson from those who have not, as stated by (Mahmudah, 2018).

The following are the results of testing the differential power index for each multiple choice question item. The differential power index for each multiple choice question item is presented in the table.

Question item no	Corrected item-total correlation	INDEX
1	0,406	Good
2	0,492	Good
3	0,435	Good
4	0,338	Good
5	0,412	Good
6	0,484	Good
7	0,501	Good
8	0,487	Good
9	0,476	Good
10	0,477	Good
11	0,500	Good
12	0,496	Good
13	0,487	Good
14	0,501	Good
15	0,454	Good
16	0,452	Good
17	0,500	Good
18	0,482	Good
19	0,500	Good
20	0,500	Good
21	0,498	Good
22	0,482	Good
23	0,498	Good
24	0,501	Good
25	0,479	Good
26	0,489	Good
27	0,501	Good
28	0,439	Good
29	0,500	Good
30	0,485	Good
31	0,501	Good
32	0,499	Good
33	0,493	Good
34	0,417	Good
35	0,483	Good
36	0,498	Good
37	0,499	Good

38	0,501	Good
39	0,429	Good
40	0,479	Good

INDEX	
0,70-1,0	Very good
0,40-0,69	Good
0,20-0,39	Enough
0,0-0,19	signs

Based on the table above, it can be seen that in the odd semester UAS questions for the Arabic language subject in the independent curriculum for class

3.Validity

Validity is often interpreted as accuracy or suitability. A measuring instrument is considered valid if the contents of the instrument adequately measure the object that should be measured and comply with certain criteria (Suzana, 2017). In this context, there is a match between measuring instruments, measurement functions, and measurement objectives. The general definition of item validity states that an item is considered valid if it significantly contributes to the total score. The score obtained on this item influences the increase or decrease in the overall score (Irawati et al., 2020). The validity of an item is declared high when the score on the item is parallel to the total score. This alignment can be measured by correlation, so that the correlation formula is used to assess the validity of the items (Abd Wahab Rosyidi & Mamlu'atul Ni'mah, 2011).

QUESTION ITEM NO	NALAI PERSON CORELLATION	SIG(2)	R.TABLE	INFORMATION
1	0,389	0,000	0,05	VALID
2	0,425	0,000	0,05	VALID
3	0,473	0,000	0,05	VALID
4	0,463	0,000	0,05	VALID
5	-0,28	0,63	0,05	INVALID
6	0,519	0,000	0,05	VALID
7	0,470	0,000	0,05	VALID
8	0,508	0,000	0,05	VALID
9	0,497	0,000	0,05	VALID
10	0,588	0,000	0,05	VALID
11	0,418	0,000	0,05	VALID
12	0,370	0,000	0,05	VALID
13	0,391	0,000	0,05	VALID
14	0,447	0,000	0,05	VALID
15	0,404	0,000	0,05	VALID
16	0,168	0,003	0,05	VALID
17	0,443	0,000	0,05	VALID
18	0,310	0,000	0,05	VALID
19	0,503	0,000	0,05	VALID
20	0,454	0,000	0,05	VALID
21	0,576	0,000	0,05	VALID

22	0,539	0,000	0,05	VALID
23	0,471	0,000	0,05	VALID
24	0,539	0,000	0,05	VALID
25	0,528	0,000	0,05	VALID
26	0,419	0,000	0,05	VALID
27	0,342	0,000	0,05	VALID
28	0,489	0,000	0,05	VALID
29	0,498	0,000	0,05	VALID
30	0,287	0,000	0,05	VALID
31	0,327	0,000	0,05	VALID
32	0,382	0,000	0,05	VALID
33	0,508	0,000	0,05	VALID
34	0,457	0,000	0,05	VALID
35	0,583	0,000	0,05	VALID
36	0,405	0,000	0,05	VALID
37	0,549	0,000	0,05	VALID
38	0,390	0,000	0,05	VALID
39	0,564	0,000	0,05	VALID
40	0,485	0,000	0,05	VALID

The validity test of the items, both multiple choice questions, was calculated using IBM SPSS 26 for Windows software, with N = 319 and a significance level of 0.05. The validity test is calculated for each type of question to determine whether or not the questions are valid for each item in each type of question. The results of calculating the validity of each multiple choice question item using IBM statistics for windows software are presented by IBM. Based on the table above, it can be seen that of the 40 questions on the Arabic language exam for the independent odd semester multiple choice curriculum that were analyzed, there were 39 (97.5%) questions with valid criteria, and 1 (2.5%) question with invalid criteria. This means that of the 40 questions, the invalid question is question number 5. The factors that influence validity are internal and external factors, internal test factors such as test directions which are prepared with unclear meaning. Meanwhile, external factors include insufficient processing time, so students become nervous and tense because they are in a hurry and can even commit cheating.

4. Reliability

Reliability comes from the term reliability, which refers to how trustworthy the measurement results are. When measurement results are carried out repeatedly on the same group of subjects and produce relatively similar values (Rudini, 2020). the measurement results are considered reliable (Wulandari et al., 2022). This applies as long as the elements measured in the subject group do not change (Huda, 2019). told the journal Validity and Reliability of a Research Instrument that a measure's reliability includes how consistent an individual's deviation scores, or z-scores, are when a test is retaken with the same or equivalent test (Zhang et al., 2014).

Reliability Statistics	
Cronbach's Alpha	N of Items
,736	40

reliability test used in this research is Coefficient alpha (Cronbach's Alpha). Reliability in this study uses a composite score reliability test which is calculated based on strata reliability, each of which is treated as a single subtest. Reliability is calculated for each type of question, then calculated using the stratified alpha formula. The results of the reliability test calculations with the help of the IBM statistics for windows software method obtained a value in Cronbach's Alpha, so the interpretation was obtained that the test instrument in the form of 40 multiple choice items was reliable because the coefficient alpha value = $0.736 > 0.70$.

5. Distracting power

When discussing objective tests in the form of multiple choice items, each item issued in the learning outcomes test is equipped with several possible answers, or what is often known as options or alternatives. Options. . Each distractor can be said to be working if at least 5% of the number of participants is selected. The following is the distracting power of the 2023/2024 odd semester final exam questions for class X Arabic at SMA 4 Muhammadiyah Jakarta:

INDEX	TOTAL
ACCEPTED	189(94,5%)
REVISION	11(5,5%)
REJECT	0(0%)

CRITERIA	
D>5%	ACCEPTED
5%>D>0	REVISION
0	REJECTED

Based on the table above, it can be seen that the odd semester UAS questions for the Arabic language subject in the independent curriculum for class .5%) options with a revised index and 0(0%) options with an index rejected.

6.PAP

The benchmark approach (PAP) or often referred to as criterion-referenced evaluation indicates the extent to which students achieve predetermined standards of success criteria, and the grades obtained do not depend on the achievements of other students. PAP, or often referred to as the ideal approach, describes a situation where students are expected to be able to answer all questions correctly and demonstrate full mastery of all the skills being tested (Arulampalam Kunaraj, P. Chelvanathan, Ahmad AA Bakar, 2023). The following are student scores that have been changed using the benchmark approach:

TOTAL PAP VALUE	
A	129
B	90
C	80
D	20
AND	0
TOTAL	319

NO	MARK			PREDICAT E
1	0	UNTIL	23	AND

2	24	UNTIL	37	D
3	38	UNTIL	52	C
4	53	UNTIL	67	B
5	68	UNTIL	90	A

In PAP, the basic assumption is that almost all students can learn, but the time they need can be different. (Muhammadiyah et al., 2020) The consequences of this approach include remediation. Students who do not reach the expected standards are required to involve themselves in additional learning or repeat learning activities until they reach the predetermined standard scores (Widiyono, 2021). The PAP shows that 129(40%) students with A grades, 90(28%) students with B grades, 80(25%) students with C grades, 20 (7%) students with D grades, and 0(0%) students with grades E.

7.PAN

Norm-referenced assessment is an assessment method that uses group norms or standards as a reference. In PAN, the scores obtained by a student are compared with the results of other students in their group (Asrul et al., 2014). This approach can be considered a classic method because it evaluates student learning outcomes based on comparisons with other students who took the same test. (Aliyani et al., 2021) PAN uses competitive learning principles and provides an overview of a student's position in their group, referring to the score obtained by other students in the group as a benchmark. Therefore, determining the PAN score is relative (Alfath, 2019). The following are student grades that have been changed using the norm reference assessment:

TOTAL PAN VALUE	
A	19
B	90
C	115
D	75
AND	20
TOTAL	319

NO	MARK			PREDICAT E
1	0	UNTIL	36	AND
2	37	UNTIL	54	D
3	55	UNTIL	71	C
4	72	UNTIL	88	B
5	89	UNTIL	100	A

Based on the conversion of PAN scores, it can be shown that 19(6%) students had an A grade, 90(28%) students had a B grade, 115(36%) students had a C grade, 75(23%) students had a D grade, and 20 (7%) students with E grades.

CONCLUSION

Based on the results of the research and also the discussion regarding the analysis of the questions on the Mid-Semester Assessment Examination for Arabic Language Subject Class with valid criteria, and 1 (2.5%) question with invalid criteria. Based on the reliability aspect, it shows a value of 0.736. Therefore, this item has high reliability. Meanwhile, in terms of level of difficulty, it shows that 5(12.5%) questions have a difficult index, 3(7.5%) questions have an easy index, and 32(80%) questions have a medium index. In terms of discriminating power, it shows that 40 (100%) of the 40 multiple choice questions had a good index. in terms of distracting power, it shows that 189(94.5%) with the option accepted and 11(5.5%) with the option revised. Student grades have been converted to PAP and show that 129(40%) students with A grades, 90(28%) students with B grades, 80(25%) students with C grades, 20 (7%) students with D grades, and 0(0%) students with a grade of E. Based on the conversion of PAN scores it can be shown that 19(6%) students with a grade of A, 90(28%) students with a grade of B, 115(36%) students with a grade of C, 75 (23%) students with a D grade, and 20(7%) students with an E grade. Based on the results of the research and analysis carried out, a conclusion can be obtained, namely the analysis of the final semester exam questions for class X Arabic at Sma 4 Muhammadiyah Jakarta with using validity, reliability, difficulty index, discriminating power, pap, and pan. makes it easy for teachers to evaluate learning and find out students' achievement or knowledge of the material. A good measuring tool will certainly have a positive impact on learning. Therefore, it is necessary to know the quality of the questions given to students to increase their cognitive level to suit learning objectives.

REFERENCE

- Abd Wahab Rosyidi & Mamlu'atul Ni'mah. (2011). Understanding the Basic Concepts of Learning Arabic. In *Journal of Chemical Information and Modeling*.
- Alfarisa, F., & Purnama, D. N. (2019). *Analysis of Final Semester Examination Questions for High School Economics Subjects Using the RASCH Model*. 11(2).
- Alfath, K. (2019). Assessment Results Processing Technique: Processing Technique Using the Norm Reference Approach (Pan) and the Benchmark Reference Approach (Pap). *Al-Manar*, 8(1), 1–28. <https://doi.org/10.36668/jal.v8i1.105>
- Aliyani, D. N. A., Dayanti, F., & Mukaffa, Z. (2021). Implementation of the Normative Reference Assessment Approach (PAN) and the Benchmark Assessment Approach (PAP) in Evaluating Student Learning Outcomes (Case Study of Class XI Students at Sma Gema 45 Surabaya). *AT-TA'LIM: Islamic Education Information Media*, 20(2), 183–191. <https://doi.org/10.29300/atmipi.v20.i2.4953.183>
- Arulampalam Kunaraj, P. Chelvanathan, Ahmad AA Bakar, I. Y. (2023). Analysis Of Classrooms Assessment: Remedial, Enrichment, Benchmark Approach (Pap) And Normative Reference Approach (PAN). *Journal of Engineering Research*, 09, 3216–3230.
- Asrul, Ananda, R., & Rosinta. (2014). Teaching Evaluation. In *Ciptapustaka Media*.
- Evi Nurus Suroiyah, & Dewi Anisatuz Zakiyah. (2021). Development of Arabic in Indonesia. *Muhadasah: Journal of Arabic Language Education*, 3(1), 60–69. <https://doi.org/10.51339/muhad.v3i1.302>

- Fathiyah, S. F. Al. (2019). Analysis of Arabic Language Lesson Question Items at MA Roudlotul Ulum Pagak Malang. *Tarbiyatuna: Journal of Scientific Education*, 4(1), 77.
- Fikriyah, N. (2021). Analysis of Mid-Semester Test Question Items for Class VII Arabic Subject, Odd Semester, SMP Muhammadiyah 1 Yogyakarta, 2019/2020 Academic Year. *Maharaat: Journal of Arabic Language Education*, 3(2), 128-140. <https://doi.org/10.18196/mht.v3i2.10501>
- Fitriyanawati, M. (2015). The Role of Question Item Analysis to Improve the Quality of Question Items, Teacher Competence and Student Learning Outcomes. *Proceedings of the National Seminar on PGSD UMS & HDPGSDI Java Region Education*, 5(3), 282-295.
- Genetik, K. S., & Khairurrijal, A. (n.d.). Artikel. *Ageing International*, 2(3), 12-13. <https://doi.org/10.1007/BF03187648>
- Halik, A. S., Mania, S., & Nur, F. (2019). Analysis of Final School Examination Questions (UAS) for Mathematics Subjects in the 2015/2016 Academic Year at SMP Negeri 36 Makassar. *Al Asma: Journal of Islamic Education*, 1(1), 11. <https://doi.org/10.24252/asma.v1i1.11249>
- Hamimi, L., Zamharirah, R., & Rusydy, R. (2020). Detailed Analysis of Mathematical Test Questions Class VII Odd Semester Academic Year 2017/2018. *Mathema: Journal of Mathematics Education*, 2(1), 57. <https://doi.org/10.33365/jm.v2i1.459>
- Hendrayani, A. E. (2016). Analysis of Even Semester Midterm Test Question Items for Class IV Indonesian Language Subjects at Gugus Antasari State Elementary School, West Tegal District, 2015/2016 Academic Year. *Joyful Learning Journal*, 2(2).
- Hidayati, K., & Nisa', Z. E. K. (2023). Analysis of Final Mathematics Assessment Question Items. *AKSIOMA: Journal of the Mathematics Education Study Program*, 12(3), 3516. <https://doi.org/10.24127/ajpm.v12i3.7575>
- Huda, N. (2019). *Curriculum Development Approaches. II* (September), 175-197.
- Ida, F. F., & Musyarofah, A. (2021). Validity and Reliability in Question Item Analysis. *Al-Mu'Arrib: Journal of Arabic Education*, 1(1), 34-44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Irawati, R., Ekawati, E. Y., & Budiawanti, S. (2020). Analysis of odd semester final exam questions using the Anbuso program at SMA Negeri 1 Boyolali for the 2019/2020 academic year. *Journal of Physics Materials and Learning*, 10(1), 11. <https://doi.org/10.20961/jmpf.v10i1.42084>
- Jufni, M., & Ibrahim, S. (2015). Teacher Pai's Creativity in Developing Teaching Materials at Madrasah Aliyah Jeumala Amal Lueng Putu. *Journal of Educational Administration: Unsyiah Postgraduate Program*, 3(4), 64-73.
- Kurniawan, R. Y., Fiky Prakoso, A., Hakim, L., Mustika Dewi, R., & Widayanti, I. (2017). Providing Question Item Analysis Training for Teachers in Jombang Regency: Effective? *Journal of Civil Society Empowerment (JPMM)*, 1(2), 179-193. <https://doi.org/10.21009/jpmm.001.2.03>
- Mahmudah, S. (2018). Arabic Learning Media. *An Nabighoh Journal of Arabic Language Education and Learning*, 20(01), 129. <https://doi.org/10.32332/an-nabighoh.v20i01.1131>
- Makhisoh, I., & Suroyya, E. N. (2022). Detailed Analysis of Arabic Language Madarasah Test Questions According to Bloom's Taxonomy. *Muhadassah*, 4(1), 40-52.
- Mania, S., Fitriani, F., Majid, A. F., Ichiana, N. N., & Abrar, A. I. P. (2020). Analysis of School Final Exam Question Items. *Al Asma: Journal of Islamic Education*, 2(2), 274. <https://doi.org/10.24252/asma.v2i2.16569>
- Marsela Yulianti, Divana Leli Anggraini, Siti Nurfaizah, & Anjani Putri Belawati Pandiangan. (2022). The Role of Teachers in Developing an Independent Curriculum. *Journal of Education and Social Sciences*, 1(3), 290-298. <https://doi.org/10.58540/jipsi.v1i3.53>
- Maulana, D., & Sanusi, A. (2020). Analysis of Arabic Language Question Items in the 2017-

- 2018 Joint Regional Madrasah (Uambd) Madrasah Ibtidaiyah Examination. *Ta'lim Al-'Arabiyyah: Journal of Arabic & Arabic Language Education*, 4(1), 12–24. <https://doi.org/10.15575/jpba.v4i1.8054>
- Muhammadiyah, U., Hamka, P., Latif, A., Zamzam, F., Kuliah, M., Phuang Phaka Masena and prayer Nanthiyakul, Fabiana Meijon Fadul, Muhammadiyah, U., Hamka, P., Kaddi, S.M., Kuliah, M., Muhammadiyah, U., Hamka, P., Studi, K.H., Hasim, R., Sofyan, A. M., ... Hasanuddin, U. (2020). Muhammadiyah Universities prof. dr. hamka. *Journal of Cultural Sciences*, 4(2), 2020. <http://journal2.um.ac.id/index.php/JPN/article/view/6091/3117%0Ahttps://www.ptonline.com/articles/how-to-get-better-mfi-results>
- Muluki, A. (2020). Analysis of the Quality of Odd Semester Test Items for Class IV Science Subjects Mi Radhiatul Adawiyah. *Primary School Scientific Journal*, 4(1), 86. <https://doi.org/10.23887/jisd.v4i1.23335>
- Nayla Amalia, A., & Widayati, A. (2012). Analysis of Quality Control Test Items for Class XII High School Economics Accounting Subjects in the City of Yogyakarta. *Indonesian Journal of Accounting Education*, X(1), 1–26.
- Nurjanah, & Marliansih, N. (2015). Analysis of Multiple Choice Questions from the Linguistic Aspect. *Journal of Educational Science Factors*, II(1), 69–78.
- Pardimin, P., Widodo, S. A., & Purwaningsih, I. E. (2017). Analysis of Mathematical Problem Solving Test Items. *ACADEMIC DISCOURSE: Educational Scientific Magazine*, 1(1), 69–76. <https://doi.org/10.30738/wa.v1i1.1084>
- Putri, N. A., Z, F., & Fauza, N. (2023). Validity and Reliability of Question Items Based on Critical Thinking Ability. *Journal of Physics Education*, 12(1), 28. <https://doi.org/10.24114/jpf.v12i1.42833>
- Rachmawati, M. (2021). Formation of an Arabic Language Environment Based on "Bi'Ah Lughowiyah" for PBA (Arabic Language Education) Uhamka Jakarta Students (Strategy and Implementation). *Al-Fakkar*, 2(2), 62–81. <https://doi.org/10.52166/alf.v2i2.2632>
- Rudini, M. (2020). Effectiveness of Analysis of Mathematics Subject Questions for Class IV Students in Improving Teacher Quality at SDN Sabang. *Scientific Writing: Research Journal*, 2(1), 17–27. https://ojs.umada.ac.id/index.php/Tolis_Ilmiah/article/view/90
- Sappaile, Baso, I. (2007). Concept of Educational Research Instruments. *Journal of Education and Culture*, No. 066(May 2007).
- Saputra, H. D., Purwanto, W., Setiawan, D., Fernandez, D., & Putra, R. (2022). Student Learning Outcomes: Test Item Analysis. *Education: Journal of Education*, 20(1), 15–27. <https://doi.org/10.31571/edukasi.v20i1.3432>
- Sari, V. N. I., Utomo, A. P. Y., & Sumarwati. (2022). Quality of Indonesian Language Questions at SMP Muhammadiyah 1 Pontianak: Analysis of Vina Question Items. *Journal of Indonesian Language and Literature Education*, 11(2), 112–119. <https://journal.unnes.ac.id/sju/index.php/jpbsi/article/view/24018>
- Statistics, M. K. (2023). *EDUCARE: Journal of Educational Sciences*. 02(01), 16–24.
- Suzana, A. (2017). Analysis of the Level of Difficulty and Different Power of End-of-Year Assessment Questions for Class X Mathematics at SMA Negeri 1 Purbalingga. *MathGram Mathematics*, 2(2), 1–8.
- Tarmizi, P., Setiono, P., Amaliyah, Y., & Agrian, A. (2021). Analysis of Multiple Choice Questions on the Theme Healthy is Important for Class V SD Negeri 04 Bengkulu City. *ELSE (Elementary School Education Journal): Journal of Elementary School Education and Learning*, 4(2), 124. <https://doi.org/10.30651/else.v4i2.7090>
- Tilaar, A. L. F., & Hasriyanti, H. (2019). Analysis of Odd Semester Question Items in Mathematics Subjects in Junior High Schools. *Indonesian Journal of Psychological and*

- Educational Measurement (JP3I)*, 8(1), 57–68.
<https://doi.org/10.15408/jp3i.v8i1.13068>
- Utomo, B. (2019). Analysis of the Content Validity of Question Items as an Effort to Improve the Quality of Learning in Madrasas Based on Islamic Values. *Journal of Mathematics Education (Holy)*, 1(2). <https://doi.org/10.21043/jpm.v1i2.4883>
- Verawati, Y., Siskawati, F. S., & Susilaningtyas, T. (2023). Analysis of Final Semester Examination (UAS) Questions for Mathematics Subjects in the 2020/2021 Academic Year Class VII Islamic Middle School at Tanwir, Ledokombo District, Jember Regency. *Education Window Journal*, 3(01), 114–121.
<https://doi.org/10.57008/jjp.v3i01.422>
- Widiyono, A. (2021). Application of ZD SOft Screen Recorder on Learning Results for PAP & PAN Material in the New Normal Era. *Proceedings of the National Seminar on Educational Policy Direction and Research Studies in the New Normal Era*, 3(1), 269–280.
- Wulandari, T., Ramli*, M., & Muzzazinah, M. (2022). Analysis of Dynamic Assessment Question Items to Measure Students' Understanding of Plant Classification Concepts. *Indonesian Journal of Science Education*, 10(1), 191–201.
<https://doi.org/10.24815/jpsi.v10i1.22082>
- Zhang, H. M., Peh, L. S., & Wang, Y. H. (2014). Servo motor control system and method of auto-detection of types of servo motors. *Applied Mechanics and Materials*, 496–500(1), 1510–1515. <https://doi.org/10.4028/www.scientific.net/AMM.496-500.1510>
-

Copyright Holder :

© M Taufiq Arkan, Ahmad Rizki Nugrahwan (2024).

First Publication Right :

© Bulletin of Science Education

This article is under:

CC BY SA