

## EVALUATING THE QUALITY OF MAWARIS AKM (MINIMUM COMPETENCY ANALYSIS (AKM) ITEMS USING RASCH MINISTEP MODELING: LITERACY AND NUMERACY

<sup>1</sup>Sofia Maulidia, <sup>2</sup>Hilmawan Putranta,

<sup>1,2</sup>, UIN Sunan Kalijaga, Yogyakarta

email: [22104010054@student.uin-suka.ac.id](mailto:22104010054@student.uin-suka.ac.id)

**Abstract :** *Item quality analysis is crucial for enhancing student capabilities and the overall quality of educational practices. The objective of this study is to analyze and delineate the quality of items associated with the Minimum Competency Assessment of Mawaris material utilizing the Rasch ministep model, thereby providing recommendations aimed at improving assessment questions to better achieve educational objectives. This research uses a quantitative descriptive approach with an evaluative focus. Data collection was executed through documentation techniques, specifically using an instrument comprising 10 AKM questions pertaining to Mawaris material, distributed to Islamic Religious Education students at UIN Sunan Kalijaga during the 2024/2025 academic year. The instrument encompassed 5 literacy questions and 5 numeracy questions. The findings indicate that out of all the questions evaluated, only 8 attained the standards of validity. Based on item estimation, item reliability was calculated at 0.81, which is regarded as exhibiting good reliability, whereas the person estimation yielded a reliability score of 0.00, which is considered weak. The analysis revealed that 7 items functioned effectively for distractor efficacy, while 3 items failed to perform adequately. Furthermore, the difficulty analysis of the evaluated items illustrates a proportional variation in the complexity of the questions.*

**Keywords:.** *Minimum Competency Assessment; Item Analysis; Ministep; Mawaris*

### Introduction

The implementation of AKM is anticipated to elevate the quality of education in Indonesia by refining the assessment methods utilized in the learning process. The outcomes of this assessment will enable educators to better understand the varying levels of student ability. The AKM material is primarily concentrated on the mastery of literacy and numeracy skills, which will be specifically measured.<sup>1</sup> The implementation of AKM is anticipated to elevate the quality of education in Indonesia by refining the assessment methods utilized in the learning process. The outcomes of this assessment will enable educators to better understand the varying levels of student ability. The AKM material is primarily concentrated on the mastery of literacy and numeracy skills, which will be specifically measured.<sup>2</sup>

Literacy encompasses not only the ability to read but also the capacity to analyze a text and comprehend the underlying meanings present within a piece of writing. In parallel, numeracy refers to the skill of interpreting data presented in numerical form.<sup>3</sup> Both of these competencies extend beyond the confines of language and mathematics education;

---

<sup>1</sup> Kementerian Pendidikan dan Kebudayaan, "AKM Dan Implikasinya Pada Pembelajaran," *Pusat Asesmen Dan Pembelajaran Badan Penelitian Dan Pengembangan Dan Perbukuan Kementerian Pendidikan Dan Kebudayaan*, 2017, 1–37.

<sup>2</sup> Dhina Cahya Rohim, "Konsep Asesmen Kompetensi Minimum Untuk Meningkatkan Kemampuan Literasi Numerasi Siswa Sekolah Dasar," *Jurnal VARIDIKA* 33, no. 1 (2021): 54–62, <https://doi.org/10.23917/varidika.v33i1.14993>.

<sup>3</sup> S Dewayani et al., *Panduan Penguatan Literasi & Numerisasi Di Sekolah* (2021).

they promote students' ability to apply these skills in the analysis of various materials. Sri Wahyuningsih, the Director of Primary Schools at the Ministry of Research and Technology, has emphasized that literacy and numeracy are crucial in enabling learners to navigate diverse scientific fields, particularly in processing and managing information, whether it is conveyed in written text or numerical figures.<sup>4</sup> This principle is also applicable to the subject of Islamic Religious Education, specifically with the theme of mawaris. The integration of literacy and numeracy skills enhances students' comprehension of mawaris, which primarily relies on text-based content. As a result, learners are encouraged to critically analyze the readings and utilize the information acquired to address problems in various contexts.

The study of mawaris holds significant importance in both academic and practical contexts due to its direct relevance to everyday life. Mawaris, derived from the plural form of al-mirats, pertains to the inheritance or transfer of assets, knowledge, property, and honor from one individual or group to another, occurring primarily as a result of death. This transfer encompasses both movable and immovable properties that are bequeathed to the surviving heirs.<sup>5</sup> Property plays a crucial role in human society, influencing social structures and serving as a means to mitigate potential conflicts arising from perceived injustices in property distribution.<sup>6</sup> Disputes often arise from individuals' insufficient comprehension of inheritance division according to Islamic law.<sup>7</sup> By acquiring knowledge in the science of mawaris, individuals can significantly reduce the likelihood of property disputes by understanding their entitlements and the proper allocation of assets as delineated in the Al-Quran.

The study of mawaris extends beyond mere textual comprehension; it necessitates enhancement through practical application to facilitate a deeper understanding of the subject matter by students. Learners should be encouraged to engage with real-world scenarios, particularly those that involve understanding and calculating inheritance issues. To effectively tackle these problems, proficiency in literacy and numeracy is essential. The Assessment Competency Measurement (AKM) established by the government is designed not only to gauge the depth of learners' competencies but also to encompass more than mere content mastery.<sup>8</sup> The outcomes of the AKM can serve as a valuable instrument for evaluating the educational process,<sup>9</sup> thereby highlighting the need for its integration across various academic disciplines.

---

<sup>4</sup> "Persiapkan Asesmen Kompetensi Minimum (AKM) Menuju Sekolah Berkualitas," Galeri Informasi, Aktivitas, Dan Transformasi Sekolah Dasar, 2020, <https://ditpsd.kemdikbud.go.id/artikel/detail/persiapkan-asesmen-kompetensi-minimum-akm-menuju-sekolah-berkualitas>.

<sup>5</sup> Masfi Sya'fiatul Ummah, "Buku Mawaris," *Sustainability (Switzerland)* 11, no. 1 (2019): 1–14.

<sup>6</sup> Kementrian Agama, "Menggapai Berkah Dengan Mawaris," *SMA Kelas XII*, 2020, 226.

<sup>7</sup> Tati Nurjanah, "Model-Model Pembelajaran Ilmu Farâ' Idh سائلنا موملوعن ارقلا او ملعة الله لوسر لاقل لاق Idh دوعسم نبا الله دبعة ض بغييس ملعلا ن او ضوبقم ورمما ذ افسائلنا موملوعن ارقلا وملعتو هاور ) اهيدى ضيقى فنم نادجي لا تمضيرقلا فى فنانثلا فالتجى فى ندنقلار هظتو ( مكاحلا," *Jurnal Penelitian Pendidikan Islam* 7, no. 2 (2019).

<sup>8</sup> Kementrian Pendidikan dan Kebudayaan, "AKM Dan Implikasinya Pada Pembelajaran."

<sup>9</sup> Atikah Suri Hasibuan, "Penerapan Konsep Asesmen Kompetensi Minimum (AKM) Dalam Meningkatkan Kemampuan Literasi Dan Numerasi Pada Siswa Di Sekolah Menengah Pertama Swasta Budi Insani," *SAJANA: Public Administration Review* 02, no. 02 (2023): 1–6.

A well-structured question is one that accurately yields information pertinent to its intended purpose, including the assessment of students' mastery of material. It is imperative to meticulously identify each assessment item prior to its administration, ensuring that the questions are effectively designed, refined, and enhanced.<sup>10</sup> Administration of tests serves as a method to evaluate an individual's proficiency following their engagement with specific content.<sup>11</sup> The quality of assessment items significantly influences students' performance, as these items produce scores that reflect the extent of comprehension regarding the material covered during instruction. Furthermore, it is equally vital to analyze questions across various interdisciplinary fields, including those related to mawaris. Consequently, educators are encouraged to advance their competencies in the formulation of assessment items. To ascertain the quality of the prepared items, it is essential to engage in comprehensive analysis activities, utilizing analytical tools such as Ministep.

The Rasch model represents a contemporary assessment theory that offers the advantage of independence between items and respondents. The model's inherent versatility empowers users with the capability to supply data across a spectrum of formats and functional configurations, thereby facilitating the rigorous analysis of a wide array of research foci. This adaptability ensures that researchers can leverage the model's capabilities to explore diverse datasets and research questions, maximizing the potential for comprehensive and nuanced findings. This model aims to establish a consistent measurement scale and curve, which facilitates the direct interpretation of student abilities. The analysis conducted through the Rasch model assessment provides comprehensive insights regarding the quality of the instruments utilized, the characteristics of the respondents, and the relationship between these two components.<sup>12</sup> The deployment of the Rasch model signifies an innovative advancement beyond classical theories that primarily focus on respondents without considering the intricacies of the items. A widely used tool for conducting Rasch model analysis is Ministep. Ministep operates as a streamlined version of Winstep, a computational tool designed to analyze score values acquired from testing instruments. Its purpose is to assess various metrics such as MNSQ Outfit, ZSTD Outfit, Point Measure Correlation, Item Reliability, Cronbach's Alpha, among others.<sup>13</sup> The Ministep software is favored for its ability to analyze Rasch models, providing users with insights into the validity, reliability, difficulty of questions, and the functionality of question distractors, as illustrated through the generated data map.

Based on the description above, this research seeks to evaluate the quality of the Minimum Competency Assessment (AKM) items pertaining to mawaris material, utilizing Ministep software. The investigation will be conducted through quantitative research,

---

<sup>10</sup> Meita Fitrianawati, "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik," *Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa 5*, no. 3 (2015): 282–95.

<sup>11</sup> Pembelajaran Di and Madrasah Berbasis, *Analisis Validitas Butir Soal Sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran Di Madrasah Berbasis Nilai-Nilai Islam*, 2019.

<sup>12</sup> Eva Dwi Kumalasari and Ihwan Mahmudi, *Analisis Pemodelan Rasch Pada Asesmen Pendidikan*, no. February (2024).

<sup>13</sup> Azizah Azizah and Sapti Wahyuningsih, "Penggunaan Model Rasch Untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria," *JUPITEK: Jurnal Pendidikan Matematika* 3, no. 1 (2020): 45–50, <https://doi.org/10.30598/jupitekvol3iss1pp45-50>.

concentrating on aspects such as validity, reliability, difficulty level, and the effectiveness of question distractors. This study is anticipated to yield valuable insights that can enhance item evaluation, thereby contributing to improved student competencies and providing educators with guidance in formulating high-quality assessment questions.

## Method

This study employs a descriptive quantitative research design, characterized by its evaluative nature. Descriptive statistics serve as the analytical method, enabling a comprehensive depiction of the collected data. This approach systematically organizes the data, facilitating easier interpretation for researchers. Furthermore, descriptive quantitative methodologies yield visual representations of data, including bar charts, pie charts, histograms, and similar graphical formats.<sup>14</sup> In selecting the sample for this research, the purposive sampling technique was employed. This technique ensures that the subjects under investigation align with the specific objectives of the research.<sup>15</sup> The aim is to include participants with relevant expertise that enhances the accuracy of the findings. The subjects of this study comprised 30 students majoring in Islamic Religious Education (PAI) at UIN Sunan Kalijaga Yogyakarta for the academic year 2024/2025.

The research instrument utilized consists of a multiple-choice questionnaire focusing on mawaris material, comprising a total of 10 questions divided equally into 5 literacy AKM questions and 5 numeracy AKM questions. The data collection technique involved administering the questionnaire through a testing approach, employing Google Forms as the medium. The testing method is employed to assess students' progress throughout the learning process<sup>16</sup> and is designed to measure respondents' abilities, attitudes, preferences, behaviors, and knowledge.<sup>17</sup> Subsequently, the data acquired will be analyzed using the Ministep tool to evaluate the validity, reliability, difficulty level, and distractor function of each question. The results will be presented through quantitative analysis to elucidate the range of values obtained from each evaluation conducted.

Table 1. Criteria for Empirical Validity

INFIT MNSQ	OUTFIT MNSQ	CORR MEASURE	Status	Desc
$0,77 \leq X \leq 1,30$	$0,5 < X < 1,5$	$0,4 < X < 0,85$	Fit	Valid
$X \geq 1,30$	$X > 1,5$	$X > 0,85$	Non Fit	Invalid
$X \leq 0,77$	$X < 0,5$	$X < 0,4$	Non Fit	Invalid

The validity of the analyzed items pertains to the validity of fit, which evaluates the appropriateness of the items in relation to the model, commonly referred to as item fit.<sup>18</sup> The assessment of validity is conducted through the analysis of INFIT MNSQ, OUTFIT

<sup>14</sup> Nurul Aziza, "Metodologi Penelitian 1 : Deskriptif Kuantitatif," *ResearchGate*, no. July (2023): 166–78.

<sup>15</sup> Ika Lenaini, "Teknik Pengambilan Sampel Purposive Dan Snowball Sampling," *HISTORIS: Jurnal Kajian, Penelitian & Pengembangan Pendidikan Sejarah* 6, no. 1 (2021): 33–39.

<sup>16</sup> Slamet Widodo et al., *Metodologi Penelitian*, in *Cv Science Techno Direct* (2023).

<sup>17</sup> M.Pd.I. Mukhamad Fathoni, *Teknik Pengumpulan Data Penelitian*, in *Jurnal Keperawatan*, no. July (2019).

<sup>18</sup> Benjamin D Wright and Mark H Stone, "Measurement Essentials," *Measurement*, 1999, 205.

MNSQ, and CORRELATION MEASURES.<sup>19</sup> An item is deemed valid if at least two of the three aspects meet the established validity standards.

Table 2. Criteria for Item and Person Estimate<sup>20</sup>

Item and Person Estimate	Criteria
$< 0,67$	Weak
$0,67 - 0,80$	Fair
$0,80 - 0,90$	Good
$0,91 - 0,94$	Very Good
$> 0,94$	Excellent

Item reliability analysis is determined through the item and person separation index.

<sup>21</sup> The numerical results derived from this analysis indicate the consistency and stability of the utilized instrument. Additionally, the effectiveness of distractors for each item is assessed by examining the number of respondents who answered correctly in comparison to those who selected a distractor. A distractor is considered effective if it is chosen by a minimum of 5% of the total respondents.<sup>22</sup> The functionality of a distractor can also be evaluated based on the mean ability; an increase in mean ability indicates that the distractor is performing its intended function appropriately.

The level of difficulty of the items is analyzed by assessing the difficulty degree associated with each item. An effective question is characterized by a balanced level of difficulty (proportional)<sup>23</sup>, as excessively easy questions may not adequately challenge students to enhance their problem-solving skills.<sup>24</sup> The criteria for evaluating the difficulty level of test items include :

Table 3. Criteria for Item Difficulty Level

Score Criteria	Category
$b > 2$	Very Difficult
$1 < b \leq 2$	Difficult
$-1 \leq b \leq 1$	Moderate
$-1 > b \geq -2$	Easy
$b \geq -2$	Very Easy

## Result and Discussion

### A. Validity Test Results for AKM Items

The validity test was conducted to assess the effectiveness of the measurement instrument utilized in this study. The analysis was completed on 10 AKM items that were distributed to respondents. The validity assessment was executed using the Ministep

<sup>19</sup> Kumalasari and Mahmudi, *Analisis Pemodelan Rasch Pada Asesmen Pendidikan*.

<sup>20</sup> Kumalasari and Mahmudi, *Analisis Pemodelan Rasch Pada Asesmen Pendidikan*.

<sup>21</sup> Widodo et al., *Metodologi Penelitian*.

<sup>22</sup> Masfi Sya'fiatul Ummah, "ANALISIS KESUKARAN SOAL, DAYA PEMBEDA DAN FUNGSI DISTRAKTOR," *Sustainability (Switzerland)* 11, no. 1 (2019): 1–14.

<sup>23</sup> Eliza Pradita et al., "Analisis Tingkat Kesukaran, Daya Pembeda, Dan Fungsi Distraktor PTS Matematika SMPN Jakarta," *Original Research* 3, no. 80 (2023): 109–18.

<sup>24</sup> Arief Aulia Rahman and Cut Eva Nasryah, *Evaluasi Pembelajaran*, in *Uwais Inspirasi Indonesia* (2019).

software, specifically in the item (column) section designated for fit order. The findings from the validity test are as follows:

Table 4. Empirical Validity Test Results

Item Number	INFIT		OUTFIT		MEASURE		Desc
	MNS Q	ZSTD	MNS Q	ZSTD	COR R	EXP	
9	1.24	0.94	1.71	1.55	0.19	0.45	Invalid
6	1.17	0.95	1.34	1.18	0.19	0.36	Valid
3	1.00	0.20	1.26	0.57	0.14	0.19	Valid
7	1.13	0.86	1.07	0.39	0.29	0.39	Valid
5	0.99	0.02	0.97	0.03	0.33	0.33	Valid
8	0.90	- 0.08	0.54	- 0.52	0.36	0.23	Valid
10	0.88	- 0.38	0.65	- 0.79	0.45	0.31	Valid
4	0.87	- 0.82	0.82	- 0.88	0.52	0.40	Valid
2	0.77	- 1.35	0.68	- 1.17	0.57	0.36	Valid
1	-	-	-	-	-	-	Undefined

According to the results obtained from the empirical validity analysis utilizing the Ministep tool, eight items satisfactorily fulfilled the validity standards, while two items failed to meet these standards and remain undefined. The items that met the validity criteria correspond to question numbers 6, 3, 7, 5, 8, 10, 4, and 2, representing a percentage of 80%. The item deemed invalid was question number 9, with a percentage of 10%. Additionally, question item 1 could not be defined within the Ministep software.

#### B. Reliability Test Results for AKM Items

Reliability for each question item is assessed through two key values: item reliability (related to the questions) and person reliability (pertaining to the respondents). The reliability assessment was conducted using the Ministep software, specifically within the summary statistic section. The results of the reliability evaluation are as follows:

Table 5. Reliability Test Results Item and Person Estimate

Type of Reliability	Score	Desc
Item Reliability	0.81	Good Reliability
Person Reliability	0.00	Weak Reliability

The item reliability value presented in Table 5 is good, indicating favorable reliability criteria. A higher item reliability value reflects a greater alignment and fit of the items within the tested model. Conversely, the person reliability value reported in Table 5 is weak, suggesting a need for improvement in reliability assessment criteria.

#### C. Distractor

The analysis of distractor function was conducted to evaluate the effectiveness of each item in misleading students into selecting incorrect answers. This assessment utilized Ministep software, specifically focusing on the "fit order" section within the item column. The results from the distractor function analysis indicate that :

Image 1. Distractor Function Ministep Test Result

Item CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

ENTRY	DATA	SCORE	DATA		ABILITY	S.E.	INFT	OUTF	PTMA	
NUMBER	CODE	VALUE	COUNT	%	MEAN	P.SD	MEAN	MNSQ	MNSQ	CORR.
9	A	0	4	13	.72	.54	.31	.6	.7	-.17
	A	0	14	47	1.02	.81	.23	1.0	1.1	-.13
	C	0	5	17	1.45	1.07	.53	1.9	1.9	.12
	B	1	7	23	1.51	1.39	.57	1.6	1.9	.19
6	B	0	2	7	-.04	.28	.28	.5	.4	-.31
	B	0	2	7	.24	.00	.00	.6	.5	-.24
	C	0	3	10	1.07	.34	.24	1.5	1.2	-.03
	D	0	2	7	2.09	.53	.53	3.7	3.7	.24
	E	1	21	70	1.30*	1.07	.24	1.1	1.1	.19
3	C	0	2	7	.62	.94	.94	1.0	1.3	-.14
	C	1	28	93	1.21	1.03	.20	1.0	1.0	.14
7	D	0	1	3	-.33	.00		.3	.3	-.27
	D	0	10	33	.88	.59	.20	1.2	1.1	-.19
	A	1	19	63	1.39	1.13	.27	1.2	1.2	.29
5	E	0	2	7	-.33	.00	.00	.4	.3	-.39
	C	0	1	3	-.33	.00		.4	.3	-.27
	E	0	2	7	.83	.00	.00	1.3	1.0	-.09
	A	0	2	7	1.56	.00	.00	2.7	2.0	.10
	D	1	23	77	1.36*	1.02	.22	.9	1.0	.33
8	d	0	3	10	.05	.26	.19	.7	.5	-.36
	B	1	27	90	1.29	1.01	.20	1.0	.9	.36
10	c	0	1	3	.24	.00		.7	.5	-.17
	A	0	1	3	.24	.00		.7	.5	-.17
	B	0	4	13	.25	.41	.24	.8	.6	-.35
	C	1	24	80	1.40	1.02	.21	1.0	.9	.45
4	b	0	11	37	.48	.62	.19	.8	.7	-.50
	D	0	1	3	.83	.00		1.0	.9	-.06
	B	1	18	60	1.60	1.03	.25	.9	.9	.52
2	a	0	2	7	-.33	.00	.00	.4	.3	-.39
	B	0	2	7	-.04	.28	.28	.5	.4	-.31
	D	0	3	10	.44	.28	.20	.8	.6	-.24
	E	0	2	7	.90	.66	.66	1.3	1.2	-.07
	C	1	21	70	1.55	.93	.21	.8	.8	.57

The efficacy of distractors can be inferred from the mean ability scores. An increase in mean ability in a sequential manner suggests that the distractor associated with a particular question item is functioning appropriately. According to the findings from the Ministep distractor function test, seven questions exhibited effective distractor functionality, specifically items 2, 3, 4, 7, 8, 9, and 10, where an increase in mean ability scores was observed. Conversely, three items—namely items 1, 5, and 6—demonstrated dysfunctional distractors.

#### D. Item Difficulty Level

An analysis of the difficulty level of test items was conducted to assess the degree of challenge presented by the questions. The difficulty of these items is determined by the competencies of respondents when addressing the AKM mawaris questions. To evaluate item difficulty, Ministep software was utilized, specifically focusing on the Wright maps within the variable section. The outcomes of the item difficulty assessment are as follows :

Table 6. Analysis of Item Difficulty Level

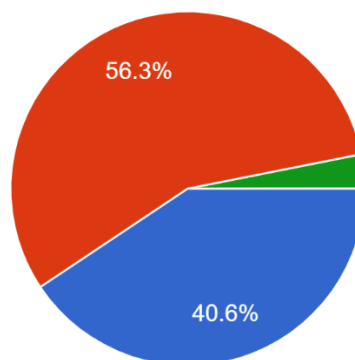
Kategori	Butir
Very Difficult	9
Difficult	-
Moderate	2, 4, 5, 6, 7, 10
Easy	3, 8
Very Easy	1

The analysis reveals a range of difficulty levels across the ten items tested. Among these, one item was categorized as very difficult, with a value greater than 2. In addition, there were six items classified within the moderate category, falling between the ranges of -1 to 1. Furthermore, two items were deemed easy, with values between -1 and -2, and one item was categorized as very easy, with a value greater than -2. Notably, no items were classified in the difficult category.

#### **E. Results of Respondents' Answers Regarding Mawaris Literacy**

Literacy questions serve to evaluate an individual's ability to comprehend written material. These questions are accompanied by a stimulus, thereby providing respondents with guidance in their responses. The results from one of the literacy question items are illustrated in a pie chart.

Image 2. Percentage of Respondents' Answers for Question Number 4



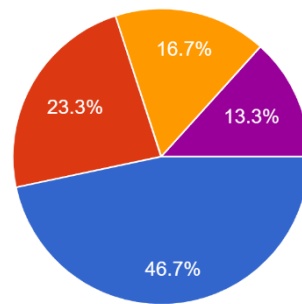
Based on the above image, Item 4 addresses the concept of distributing inheritance in accordance with Islamic law. Figure 2 presents the responses of 30 individuals who participated in this assessment. The data indicates that a majority of respondents (56.7%) selected answer B, which is identified as the correct response. However, it is noteworthy that 40.6% of respondents opted for answer A, while 16.7% chose answer D.

#### **F. Results of Respondents' Answers Regarding Mawaris Numeracy**

Numeracy questions serve to assess an individual's capacity to comprehend the various numbers and symbols utilized in the analysis of information. Each question is accompanied by a stimulus, which provides respondents with guidance and clarity in responding. The results from one particular numeracy item are displayed in the form of a pie chart.

Image 3. Percentage of Respondents' Answers for Question Number 9





Based on the above image, item 9 pertains to the calculation of inheritance distribution among members of an Islamic family. As illustrated in Figure 3, the responses of 30 participants indicate that the majority of students (46.7%) selected answer A, identified as the correct response. Nevertheless, it is noteworthy that 23.3% of the learners opted for answer B, while 16.7% chose answer C, and 13.3% selected answer E.

Validity is defined as the accuracy and precision of a test in fulfilling its measurement role.<sup>25</sup> Results from the validity assessment using Ministep software indicate that eight items satisfy the established validity standards, while two items fail to meet these standards and are considered undefined. Specifically, item number nine is deemed invalid, as it yields an OUTFIT MNSQ value of 1.71 and a CORR MEASURE value of 0.19, both of which do not conform to the requisite validity criteria for a question item. In the case of question item number one, the Ministep software does not provide a defined value due to all 30 respondents answering this specific question correctly on the AKM Mawaris assessment. Research conducted by Elviana on item analysis utilizing the Anates program revealed that among 25 tested items, 19 were deemed insignificant.<sup>26</sup> Therefore, it can be concluded that the execution of validity tests can provide researchers insights into the validity of their items, despite some results not aligning with overall validity standards.

In terms of reliability analysis, the item reliability value obtained is 0.81, categorizing it as "good," indicating that the formulated questions have met the reliability benchmarks suitable for dissemination across diverse populations. Conversely, the person reliability value is 0.00, which is classified as "weak." This suggests a likelihood that respondents with higher abilities may answer incorrectly, while those with lower abilities might answer correctly. Additionally, there is a possibility that some respondents may engage with the questions hurriedly or without due seriousness. Reliability testing conducted by Nuri et al. corroborates these findings, as they recorded an item reliability of 0.90, meeting reliability criteria, whereas the person reliability was at 0.37, indicating a lack of dependability.<sup>27</sup> Reliability denotes the consistency of measurement instruments, which

<sup>25</sup> Elviana, "Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates," *Jurnal MUDARRISUNA* 10, no. 2 (2020): 58–74.

<sup>26</sup> Elviana, "Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates."

<sup>27</sup> N S Azzahra et al., "Analisis Validitas Dan Realibilitas Kualitas Soal Pilihan Ganda Asesmen Kompetensi Minimum (AKM) Mata Pelajaran Pendidikan Agama Islam Menggunakan ...," *QuranicEdu: Journal of ...* 4, no. 1 (2024): 85–94.

remain stable even when the same instrument is utilized repeatedly. It assesses the degree to which a test can produce scores that are stable and relatively unchanged.<sup>28</sup>

The distractor analysis presented in Figure 1 identified seven items that functioned effectively, achieving a percentage of 70%, while three items were categorized as ineffective with a percentage of 30%. For instance, in item 9, the mean abilities for the options were as follows: option E at 0.72, option A at 1.02, option C at 1.45, and option B at 1.51. The sequential increase in mean ability values indicates that the distractors in item 9 are functioning appropriately. Conversely, three questions with ineffective distractors were noted in items 1, 5, and 6. In item 6, the mean abilities were recorded as follows: option A at -0.04, option B at 0.24, option C at 1.07, option D at 2.09, and option E at 1.50. The absence of a sequential increase in the mean abilities signifies that the distractors for item 6 are non-functional, as indicated by the asterisk. A distractor analysis conducted by Eliza et al<sup>29</sup> on a set of 40 items revealed that 35 items exhibited functional distractors, while the remaining items did not. Another study highlighted that the minimum threshold for a distractor to be considered functional is 5%, resulting in findings where 8 questions demonstrated good distractor functions and 17 did not.

The analysis of item difficulty utilizing the Ministep software indicated variations in the difficulty levels of the questions, comprising one item categorized as very difficult, six items as medium, two items as easy, and one item as very easy. Overall, the assessment of the AKM Mawaris questions can be regarded as effective, based on the difficulty levels determined through Ministep software analysis. Moreover, the bar chart reflecting the literacy and numeracy skills of students suggests that a majority of students demonstrated an understanding of literacy concepts within the Mawaris questions, while some students faced challenges in comprehending the numerical aspects. Consequently, there remains a need to enhance students' numeracy skills. Research by Eliza et al. indicated variations in question difficulty, with four items classified as too easy, 34 items as medium, and two items as too difficult.<sup>30</sup> The item difficulty test conducted by Nuri et al,<sup>31</sup> produced results indicating two items in the very easy category, three items in the easy category, three items in the medium category, and two items in the very difficult category. In conclusion, the research instrument or questions administered to students yielded varying levels of difficulty, reflecting the abilities of each respondent.

## **Conclusion**

The conclusion of the study, which is based on the analysis of the quality of items in the Minimum Competency Assessment (AKM) for mawaris material utilizing Ministep software, indicates several key findings. The validity test results reveal that 8 items satisfy the established validity standards, while 2 items are deemed invalid and undefined. The item reliability coefficient is measured at 0.81, reflecting a strong reliability of the questions

---

<sup>28</sup> Widodo et al., *Metodologi Penelitian*.

<sup>29</sup> Pradita et al., "Analisis Tingkat Kesukaran, Daya Pembeda, Dan Fungsi Distraktor PTS Matematika SMPN Jakarta."

<sup>30</sup> Pradita et al., "Analisis Tingkat Kesukaran, Daya Pembeda, Dan Fungsi Distraktor PTS Matematika SMPN Jakarta."

<sup>31</sup> Azzahra et al., "Analisis Validitas Dan Realibilitas Kualitas Soal Pilihan Ganda Asesmen Kompetensi Minimum (AKM) Mata Pelajaran Pendidikan Agama Islam Menggunakan ...."

formulated. However, the Person Reliability is noted at 0.00, suggesting low reliability among the respondents completing the assessment. Furthermore, the analysis identifies 7 items with dysfunctional distractors and 3 items that also exhibit dysfunctionality. This assessment is corroborated by the observed increase in the mean ability of the participants. The evaluation of item difficulty demonstrates a proportionate variation, with the distribution as follows: 1 item classified as very difficult, 6 items categorized as medium, 2 items identified as easy, and 1 item as very easy. Additionally, an examination of the pie chart representing answer results highlights a continued need for improvement in students' competencies in numeracy and literacy related to mawaris material.

## References

- Agama, Kementrian. "Menggapai Berkah Dengan Mawaris." *SMA Kelas XII*, 2020, 226.
- Aziza, Nurul. "Metodologi Penelitian 1 : Deskriptif Kuantitatif." *ResearchGate*, no. July (2023): 166–78.
- Azizah, Azizah, and Sapti Wahyuningsih. "Penggunaan Model Rasch Untuk Analisis Instrumen Tes Pada Mata Kuliah Matematika Aktuaria." *JUPITEK: Jurnal Pendidikan Matematika* 3, no. 1 (2020): 45–50. <https://doi.org/10.30598/jupitekvol3iss1pp45-50>.
- Azzahra, N S, S Sumarni, and H Putranta. "Analisis Validitas Dan Realibilitas Kualitas Soal Pilihan Ganda Asesmen Kompetensi Minimum (AKM) Mata Pelajaran Pendidikan Agama Islam Menggunakan ...." *QuranicEdu: Journal of ...* 4, no. 1 (2024): 85–94.
- Dewayani, S, P Retnaningdyah, D. Susanti, and B Antoro. *Panduan Penguatan Literasi & Numerisasi Di Sekolah*. 2021.
- Di, Pembelajaran, and Madrasah Berbasis. *Analisis Validitas Butir Soal Sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran Di Madrasah Berbasis Nilai-Nilai Islam*. 2019.
- Elviana. "Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates." *Jurnal MUDARRISUNA* 10, no. 2 (2020): 58–74.
- Fitrianawati, Meita. "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik." *Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa* 5, no. 3 (2015): 282–95.
- Galeri Informasi, Aktivitas, Dan Transformasi Sekolah Dasar. "Persiapkan Asesmen Kompetensi Minimum (AKM) Menuju Sekolah Berkualitas." 2020. <https://ditpsd.kemdikbud.go.id/artikel/detail/persiapkan-asesmen-kompetensi-minimum-akm-menuju-sekolah-berkualitas>.
- Hasibuan, Atikah Suri. "Penerapan Konsep Asesmen Kompetensi Minimum (AKM) Dalam Meningkatkan Kemampuan Literasi Dan Numerasi Pada Siswa Di Sekolah Menengah Pertama Swasta Budi Insani." *SAJJANA: Public Administration Review* 02, no. 02 (2023): 1–6.
- Kementrian Pendidikan dan Kebudayaan. "AKM Dan Implikasinya Pada Pembelajaran." *Pusat Asesmen Dan Pembelajaran Badan Penelitian Dan Pengembangan Dan Perbukuan Kementerian Pendidikan Dan Kebudayaan* *Pembelajaran Badan Penelitian Dan Pengembangan Dan Perbukuan Kementerian Pendidikan Dan Kebudayaan*, 2017, 1–37.
- Kumalasari, Eva Dwi, and Ihwan Mahmudi. *Analisis Pemodelan Rasch Pada Asesmen Pendidikan*. no. February. 2024.
- Lenaini, Ika. "Teknik Pengambilan Sampel Purposive Dan Snowball Sampling." *HISTORIS: Jurnal Kajian, Penelitian & Pengembangan Pendidikan Sejarah* 6, no. 1 (2021): 33–39.

- Mukhammad Fathoni, M.Pd.I. *Teknik Pengumpulan Data Penelitian*. In *Jurnal Keperawatan*. no. July. 2019.
- Nurjanah, Tati. “Model-Model Pembelajaran Ilmu Farâ’ Idh سائلنا دوعسم نارقلا اوملعة الله لوسر”. In *Jurnal Penelitian Pendidikan Islam* 7, no. 2 (2019).
- Pradita, Eliza, Priarti Megawanti, and Universitas Indraprasta PGRI. “Analisis Tingkat Kesukaran, Daya Pembeda, Dan Fungsi Distraktor PTS Matematika SMPN Jakarta.” *Original Research* 3, no. 80 (2023): 109–18.
- Rahman, Arief Aulia, and Cut Eva Nasryah. *Evaluasi Pembelajaran*. In *Uwais Inspirasi Indonesia*. 2019.
- Rohim, Dhina Cahya. “Konsep Asesmen Kompetensi Minimum Untuk Meningkatkan Kemampuan Literasi Numerasi Siswa Sekolah Dasar.” *Jurnal VARIDIKA* 33, no. 1 (2021): 54–62. <https://doi.org/10.23917/varidika.v33i1.14993>.
- Ummah, Masfi Sya'fiatul. “ANALISIS KESUKARAN SOAL, DAYA PEMBEDA DAN FUNGSI DISTRAKTOR.” *Sustainability (Switzerland)* 11, no. 1 (2019): 1–14.
- Ummah, Masfi Sya'fiatul. “Buku Mawaris.” *Sustainability (Switzerland)* 11, no. 1 (2019): 1–14.
- Widodo, Slamet, Festy Ladyani, La Ode Asrianto, et al. *Metodologi Penelitian*. In *Cv Science Techno Direct*. 2023.
- Wright, Benjamin D, and Mark H Stone. “Measurement Essentials.” *Measurement*, 1999, 205.