ISSN Online: 2442-5605

Editorial Office: Faculty of Teacher Training and Education, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
Phone: +62781365280889
E-mail: potensia.ftk@uin-suska.ac.id
Website: https://ejournal.uin-suska.ac.id/index.php/potensia

Validity and Reliability of the Artificial Intelligence Literacy for Prospective Islamic Education Teachers Using Rasch Model

Herlini Puspika Sari[a], Devi Arisanti[a], Aldeva Ilhami[a]

[a]Faculty of Education and Teaching, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

http://dx.doi.org/10.24014/potensia.v11i2.38308

Abstract

This study aims to validate an Artificial Intelligence (AI) literacy instrument for Islamic Education (PAI) students using Rasch analysis to evaluate its reliability, validity, item-fit statistics, and unidimensionality. The instrument was developed to measure the extent of AI literacy among prospective Islamic Education teachers. A cross-sectional survey design was employed, involving 64 students from a university in Riau, Indonesia. Participants completed an AI literacy instrument consisting of 85 items. The sampling technique used was quota sampling. Data were analyzed using the Rasch model with the aid of Winsteps software version 3.73. The results of the Rasch analysis indicated that the instrument demonstrated good reliability (α = 0.85), excellent item quality (0.94), and consistent respondent reliability (0.84). In terms of validity, the three dimensions of AI literacy were confirmed to possess unidimensional properties. The practical implication of this study lies in providing higher education institutions with a validated AI literacy instrument that can be used to assess the knowledge of future Islamic Education teachers. By employing Rasch analysis, this study contributes to enhancing the psychometric robustness of measurement tools, particularly in the context of assessing artificial intelligence literacy in Indonesia.

Keywords

Validity; reliability; Rasch model; AI literacy; islamic education.

Correspondence:
Herlini Puspika Sari, Faculty of Education and Teaching, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
E-mail: herlini.puspika.sari@uin-suska.ac.id

Introduction

Artificial intelligence (AI) technology has developed rapidly, creating a significant impact in the field of education. AI literacy has become an essential aspect that prospective educators must possess. AI literacy not only includes understanding how AI systems operate but also involves applying them to support more effective and efficient learning processes. As dependence on this technology increases, it becomes crucial for prospective Islamic Education (PAI) teachers to understand how AI can be integrated into PAI instruction. The ability to master this technology has become an urgent necessity for future educators (Fadlelmula & Qadhi, 2024). Prospective PAI teachers are being prepared to become professional educators who are adaptive to technological changes (Wilton et al., 2022). Therefore, developing AI literacy skills among PAI pre-service teachers is of great importance.

AI literacy can be defined as the ability to understand and use AI intelligently and responsibly. According to UNESCO (2024) AI literacy refers to the knowledge, skills, and values that instill an ethical understanding of human-built methods forming the foundation of AI systems. AI literacy equips pre-service teachers to face the challenges emerging from technological advancements. It also includes the ability to evaluate the social and ethical impacts that may arise from the use of such technology. By mastering AI literacy, students are expected to utilize AI productively and think critically about its potential risks (Zootzky & Pfeiffer, 2024). he framework of artificial intelligence literacy (AI literacy) fosters a more comprehensive understanding of AI by integrating three main components: knowledge, skills, and attitudes (OECD, 2025). KThese components are interrelated and form the foundation for learners to interact with AI effectively and ethically. Knowledge provides the theoretical basis of how AI functions, skills enable learners to utilize AI in various contexts, and attitudes encourage learners to apply AI responsibly and wisely.

AI technology offers various positive impacts that can facilitate teachers' work. In the educational context, AI can be used to simplify administrative tasks, accelerate the assessment process, and deliver learning materials tailored to individual students' needs. AI can also improve educational accessibility by providing diverse resources that students worldwide can access. This technology not only assists teachers in enhancing teaching efficiency but also enables them to focus on more strategic aspects of instruction, such as student interaction and social skills development (Faisal Tariq Hasan and Rafiq Rahman 2024; Rahayu 2023; Abbas 2023). However, the use of AI must be accompanied by a deep understanding of its potential social and ethical impacts

A review of literature on the development of AI literacy instruments indicates that while significant progress has been made, there remains a lack of focus on prospective PAI teachers. Kesuma & Fransen (2025) developed an AI literacy instrument for computer science students. Biagini et al. (2023), Grassini (2024), and Shi (2025) designed multidimensional literacy questionnaires for higher education contexts, while Ng et al. (2024) developed measurement instruments for secondary school students. However, existing instruments have not fully captured the dimensions of AI literacy relevant to religious education, particularly in understanding fundamental AI concepts, developing critical and creative thinking skills, and fostering attitudes that promote ethical technology use in religious education. This gap highlights the need for an AI literacy instrument specifically designed to assess AI literacy among prospective PAI teachers. Thus, developing such an instrument is a crucial endeavor.

On the other hand, the use of Rasch analysis in educational instrument development has proven to yield more accurate results compared to classical methods. Rasch analysis can identify biased items and provide specific recommendations for instrument revision (Müller, 2020;

Prasetya & Pratama, 2023). It enhances instrument quality by aligning item difficulty levels with the distribution of respondents' abilities (Metsämuuronen, 2023). This approach ensures that the instrument accurately differentiates respondents' competency levels and remains free from systematic bias. The novelty of this study lies in the application of Rasch analysis to examine the validity and reliability of an AI literacy instrument. The scope of the research includes item fit, unidimensionality, item and respondent reliability testing. The findings may serve as a reference for educational research and practices that support the achievement of Sustainable Development Goals (SDGs) through improving the quality of PAI pre-service teachers.

This research is essential to support curriculum development for Islamic Religious Education in higher education. With the increasing advancement of AI technology, universities must ensure that their curricula train PAI pre-service teachers to be adaptive to technological change (Sari et al., 2025). The curriculum must be aligned with contemporary needs to ensure that future teachers can effectively integrate technology into their teaching practices (Zhang & Zhang, 2024). This study is expected to make a significant contribution by developing an instrument that can be used to measure AI literacy among PAI pre-service teachers. The purpose of this study is to measure the validity and reliability of an AI literacy instrument specifically designed for PAI students. The instrument will be tested using Rasch analysis to ensure it possesses good quality in measuring AI literacy among PAI pre-service teachers.

Method

This study employed a cross-sectional survey design, focusing on collecting information from the research sample through a questionnaire. The instrument used was the AI literacy questionnaire, which followed the AI literacy framework (OECD, 2025). The instrument consisted of three dimensions knowledge, cognitive skills, and attitudes with a total of 85 items, all of which used a Likert scale. The respondents comprised 64 students majoring in Islamic Religious Education (PAI) at universities across the Riau Province. The instrument underwent content validation by three expert validators and was pilot tested to assess item readability, focusing on textual errors, typos, and image clarity.

Table 1. AI Literacy Framework

Dimension	Indicator	Items
Knowledge	The Nature of AI	5
	AI Reflects Human Choices and Perspectives	5
	AI Reshapes Human Work and Roles	5
	Capabilities and Limitations of AI	5
	The Role of AI for Society	5
Cognitive Skills	Critical Thinking	5
	Creative Thinking	5
	Self and social Awareness	5
	Collaboration	5
	Communication	5

	Problem solving	5
	Computational Thinking	5
Attitude	Responsibility	5
	Innovative	5
	Flexible	5
	Empathetic	5
	Curiosity	5

The respondents in this study were PAI pre-service teachers. A total of 64 students participated in this survey, and their demographic characteristics are presented in the table below. This number is sufficient for instrument validation using the Rasch analysis approach.

Table 2. Demographics of Respondents

Aspect		Frequency	Percentage
Gender	Male	201	29.9%
	Female	472	70.1%
Semester	1	151	22.4%
	3	166	24.7%
	5	89	13.2%
	7	256	38.0%
	8	3	0.4%
	9	4	0.6%
	11	4	0.6%
University	UIN Sultan Syarif Kasim Riau	224	33.3%
	IAIN Datuk Laksemana Bengkalis	145	21.5%
	STAI Al Azhar pekanbaru	58	8.6%
	IAI Rokan	60	8.9%
	IAI diniyah pekanbaru	21	3.1%
	STAI Nurul Falah Air Molek	62	9.2%
	STAI Sulthan Syarif Hasyim Siak Sri Indrapura Riau	30	4.5%
	Universitas Islam Kuantan Singingi	34	5.1%
	Universitas Islam Riau	25	3.7%
	IAI Dar Aswaja	7	1.0%
	Institut keislaman tuah negeri	7	1.0%

Data were collected from student answer sheets and subsequently tabulated in Microsoft Excel. The data were then coded and extracted into a raw data (.prn) format for analysis using

WINSTEPS software (version 3.73), which employs the Rasch measurement model. The psychometric properties of the instrument were evaluated against six key parameters: validity, reliability, unidimensionality, item fit analysis, partial credit rating, and a variable map.

To ascertain the suitability of each item for measuring AI literacy, three fit statistics were examined: Outfit Mean Square (MNSQ), Outfit Z-Standard (ZSTD), and Point-Measure Correlation (PTMEASURE Corr.), as detailed in Table 2. Any item that failed to meet the established criteria for two or more of these parameters was removed from the instrument. The criteria for determining instrument reliability followed the rubric presented in Table 4.

Table 3. Item fit

Criteria	Item Fit Value
Outfit mean square (MNSQ)	0,5 < MNSQ < 1,5
Outfit Z-standard (ZSTD)	-2,0 < ZSTD < +2,0
Point measure correlation	0,4< PTMeasure Corr. <0,85

(Bond et al., 2021)

Table 4. Reliability Criteria

Score	Description
<0,67	Weak
0,67-0,80	Sufficient
0,80-0,90	High
0,91-094	Very High
>0,94	Excellent

(Sumintono & Widhiarso, 2015)

Results and Discussion

The data on the knowledge aspect of student AI literacy were analyzed using the Rasch model, a modern test theory approach. This analysis yielded a comprehensive overview of the instrument's psychometric properties, including: (a) item fit, based on Mean Square, Z-Standard, and Point-Measure Correlation statistics; (b) unidimensionality; (c) item reliability; (d) person reliability; (e) partial credit rating scale function; and (f) the distribution of respondent ability and item difficulty for the AI literacy knowledge construct, visualized on a Wright Map. The detailed item statistics for Mean Square, Z-Standard, and Point-Measure Correlation are presented in Table 2.

Item Fit

Item fit analysis is conducted to ensure that each item functions appropriately within the context of the measurement objective and does not introduce confusion or misunderstanding for the respondents. The evaluation of item fit focuses on three key statistical values: Outfit Mean Square

(OUTFIT MNSQ), Outfit Z-Standard (OUTFIT ZSTD), and Point-Measure Correlation (PT-MEASURE CORR.). An item is determined to have a good fit if it aligns with the established criteria for these statistics (Sumintono & Widhiarso, 2015). The specific criteria applied to each of these three values are detailed in Table 5.

Table 5. Item Fit

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
1	AI operates using algorithms that combine step-by-step procedures and statistical inferences to process data.	0.58	-2.42	0.22	P1	Invalid
2	AI can provide recommendations based on available data, but the final decision still requires human judgment.	0.93	-0.33	0.3	P2	valid
3	AI is capable of processing data quickly and efficiently to produce useful information in education.	0.55	-2.69	0.49	P3	valid
4	AI can replace the human role in deeply understanding social and cultural contexts.	1.32	1.94	0.19	P4	valid
5	AI has a high capability to understand religious values.	1.02	0.2	0.33	S5	valid
6	AI reflects human choices in system design and training, including decisions about the data used.	1.29	1.45	0.05	P6	valid
7	AI can exhibit biases present in the training data, which affects how the system functions.	0.57	-2.84	0.4	P7	valid
8	AI is designed by considering human perspectives.	1.76	3.13	-0.21	P8	Invalid
9	AI can always perfectly reflect religious values.	1.07	0.49	0.34	P9	valid
10	Decisions made by AI systems are always in accordance with social or religious perspectives	0.78	-1.53	0.34	P10	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
11	AI can help automate administrative tasks in education.	0.63	-2.05	0.41	P11	Invalid
12	The integration of AI in education can increase learning efficiency	0.59	-2.3	0.29	P12	Invalid
13	AI can help accelerate the learning process by providing customized materials	0.59	-2.5	0.35	P13	Invalid
14	AI can fully replace the role of a teacher in teaching the moral and ethical principles of Islam.	1.79	3.29	0.44	P14	Invalid
15	When using AI, teachers do not have to be involved in assessing and guiding students according to religious teachings.	2.38	5.02	0.25	P15	Invalid
16	AI is excellent at processing large data and automating routine tasks that require high efficiency	1.75	3.23	0.03	P16	Invalid
17	AI can provide sophisticated analysis based on the data it is given	0.76	-1.23	0.43	P17	valid
18	AI has the ability to process information quickly and accurately	1.68	2.86	-0.22	P18	Invalid
19	AI cannot replace the moral or spiritual judgment required in the context of Islamic religious teachings	1.71	2.96	0.48	P19	Invalid
20	AI has no limitations in understanding knowledge, including Islamic religious teachings, because it already has valid and complete data sources	1.28	1.79	0.09	P20	valid
21	AI is able to provide appropriate learning materials so that it has the potential to increase access to education	0.91	-0.4	0.12	P21	valid
22	The use of AI in education can support the effectiveness of Islamic religious learning	0.97	-0.08	0.23	P22	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
23	AI can increase efficiency in the religious learning process because it provides fast results and according to needs	1.51	2.32	-0.18	P23	Invalid
24	The use of AI in Society must always consider its impact on social justice and diversity.	0.67	-1.89	0.5	P24	valid
25	The use of AI in religious education can lead to misuse of materials that are not in accordance with religious principles.	1.92	3.64	0.04	P25	Invalid
26	I was able to analyze the results provided by the AI system and evaluate whether they fit the context of Islamic religious teachings.	0.53	-2.78	0.51	K1	valid
27	I was able to identify biases in the outputs generated by AI and examine their truth and relevance in the context of Islamic religious education	0.66	-1.83	0.29	K2	valid
28	I have always used critical thinking to assess whether AI provides information that is in accordance with Islamic moral and ethical values.	1.16	0.83	0.34	K3	valid
29	I feel there is no need to double-check the AI-generated material because it already has a valid data source	1.22	1.24	0.57	K4	valid
30	I find it difficult to evaluate the truth of the results of the AI system	0.89	-0.72	0.27	K5	valid
31	I can use AI to help develop creative ideas in teaching Islamic religious material to students.	0.83	-0.86	0.4	K6	valid
32	I feel that AI can provide new inspiration in designing more	0.4	-3.96	0.62	K7	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
	interesting and relevant Islamic religious learning methods.					
33	I can use AI to create innovative teaching materials, such as exam questions or presentations, that fit the learning material.	0.74	-1.41	0.47	K8	valid
34	I don't feel the need to use AI in developing creative ideas for PAI teaching in the classroom, because creativity is better developed manually.	0.85	-0.95	0.46	K9	valid
35	I often find using AI to generate PAI learning strategy ideas uninteresting and unnecessary.	0.82	-1.24	0.58	K10	valid
36	I can identify problems in learning and can be solved using AI	1.59	2.56	0.17	K11	Invalid
37	I am able to use logic and systematic step-by-step in solving problems related to learning using AI	0.42	-3.61	0.4	K12	valid
38	I can formulate the problem I want to solve so that it can be understood by the AI system to obtain a relevant solution (e.g. creating commands/prompts in gpt chat)	0.59	-2.31	0.65	K13	valid
39	I find it difficult to use AI to solve the problem I want to solve	0.74	-1.83	0.49	K14	valid
40	I'm not sure that AI can help solve the learning problems I'm finding effectively	0.89	-0.69	0.38	K15	valid
41	I realized that AI is affecting my view of information related to Islam and education	1.27	1.46	0	K16	valid
42	I am able to understand the social impact of the use of AI in education, including how it can affect students	0.99	0	0.17	K17	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
43	I am aware that the use of AI can reinforce or weaken certain social values in PAI learning because it depends a lot on how the technology is used.	1.42	1.89	0.21	K18	valid
44	I don't always consider the social impact of using AI in learning because AI is just a tool	1.58	3.36	0.14	K19	Invalid
45	I find it difficult to consider the impact of using AI on students.	0.92	-0.48	0.26	K20	valid
46	I can work with AI in designing Islamic teaching materials that suit the needs and characteristics of students.	0.85	-0.71	0.27	K21	valid
47	I feel that the collaboration between me and AI technology can enrich the learning process of Islam by providing faster and more effective solutions.	0.83	-0.83	0.48	K22	valid
48	I am able to give clear commands to the AI to produce appropriate learning materials	1.12	0.66	0.36	K23	valid
49	I feel that collaboration with AI in PAI learning often does not make a meaningful contribution	0.85	-1.01	0.4	K24	valid
50	I don't feel the need to work with AI in developing Islamic teaching materials, because I prefer to do it independently.	0.84	-1.05	0.67	K25	valid
51	I was able to explain how AI works in the context of teaching Islam to fellow educators	0.72	-1.47	0.22	K26	valid
52	I was able to communicate the benefits and limitations of AI in supporting PAI learning in a way that was easy for others to understand.	0.74	-1.37	0.25	K27	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
53	I feel confident in conveying the results or information generated by AI related to PAI learning materials to students clearly and precisely.	1.83	3.52	-0.17	K28	Invalid
54	I find it difficult to explain AI working in the context of PAI due to a lack of understanding of how it works.	0.95	-0.32	0.25	K29	valid
55	I often feel that I don't need to explain how AI is used in PAI learning to students, because I feel like they already understand the technology.	0.84	-1.08	0.47	K30	valid
56	I can use AI to help solve problems that arise in PAI learning such as creating HOTS skill-based exam questions	0.83	-0.82	0.36	K31	valid
57	I feel that AI can provide effective solutions in overcoming challenges in PAI learning	1.81	3.42	-0.05	K32	Invalid
58	I can design an AI-based solution to improve the quality of Islamic religious teaching in the classroom	1.05	0.31	0.25	K33	valid
59	I feel that the use of AI cannot help in solving the problems I face in PAI learning	1.78	4.05	-0.21	K34	Invalid
60	I often feel that using AI for problem-solving in PAI learning is inefficient and ineffective.	0.69	-2.3	0.53	K35	valid
61	I feel a responsibility to convey to students that the use of AI does not violate ethical principles	1.03	0.25	0.44	S1	valid
62	I ensure that the output of AI in PAI learning always considers its impact on society and religious values.	0.72	-1.58	0.52	S2	valid
63	I recognize the importance of acting carefully when using AI	1.05	0.31	0.44	S3	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
64	I often assume that AI is a neutral technology so it doesn't need to be ethically considered in its use	1.48	2.83	0.29	S4	Invalid
65	I feel like I don't need to worry too much about social and moral responsibility when using AI in PAI learning	1	0.05	0.56	S5	valid
66	I have a high curiosity to learn more AI to use in PAI learning	1.1	0.54	0.4	S6	valid
67	I am excited to explore different applications of AI to see how this technology can help me in preparing for more effective PAI learning.	0.75	-1.31	0.45	S7	valid
68	I strive to understand how AI works in the context of PAI learning in depth.	0.69	-1.65	0.45	S8	valid
69	I feel less interested in exploring further the use of AI	0.97	-0.16	0.58	S9	valid
70	I rarely feel interested in learning more about AI technology because I feel irrelevant to PAI learning	0.94	-0.32	0.59	S10	valid
71	I love finding new ways to integrate AI in PAI learning	1.24	1.17	0.37	S11	valid
72	I believe that AI can be an innovative tool in supporting more interesting and relevant Islamic religious learning for today's generations.	1.02	0.15	0.22	S12	valid
73	I am ready to adapt AI technology in designing PAI learning methods that are more creative and in line with the times.	0.71	-1.58	0.47	S13	valid
74	I feel like the use of AI in PAI learning won't bring much significant change	0.87	-0.85	0.58	S14	valid
75	I feel there is no need to use AI in PAI learning because I believe	0.91	-0.58	0.64	S15	valid

No	Statements	MNSQ	ZSTD	PT-Measure (CORR)	Item	Information
	more in the conventional approach that already exists					
76	I find it easy to adapt to changes and technological advancements, including in the use of AI	1.38	1.73	-0.03	S16	valid
77	I was able to quickly adjust to the use of AI which can help improve competence as a prospective teacher	0.85	-0.73	0.19	S17	valid
78	I am ready to adapt the teaching methods in PAI by integrating AI technology.	0.81	-0.95	0.33	S18	valid
79	I often find it difficult to adapt to new technologies, including AI	0.95	-0.26	0.42	S19	valid
80	I find it awkward to use new technologies in PAI teaching including AI technology	0.77	-1.65	0.6	S20	valid
81	I strive to always consider the social and emotional impact of using AI in PAI	0.62	-2.1	0.53	S21	valid
82	I realized the importance of understanding students' perspectives when using AI in PAI learning so that this technology can be used for good	0.8	-1.05	0.45	S22	valid
83	I strive to leverage AI that supports students' social and emotional development	0.77	-1.2	0.34	S23	valid
84	The use of AI is not important to be feared to affect students socially and emotionally	0.93	-0.37	0.63	S24	valid
85	AI technology is already designed for good so there is no need to supervise students in using it	1.55	2.72	0.5	S25	valid

Table 5 displays the measurement results for each item, including its Mean Square (MNSQ), Z-Standard (ZSTD), and Point-Measure Correlation (Pt-Measure Corr.) values. The criteria for acceptable fit were defined by a value between -1.5 and 1.5, a Z-Standard score in the

range of -0.2 < ZSTD < +2.0, and a Point-Measure Correlation between 0.4 and 0.85 (Boone, 2016). However, the analysis revealed that 18 items did not meet these criteria and were thus considered to have a poor fit (misfit) and deemed unacceptable. These items were: P1, P8, P11, P12, P13, P14, P15, P16, P18, P19, P23, P25, K11, K19, K28, K32, K34, and S4. Conversely, the remaining 67 items satisfied the criteria and were retained for the final instrument.

The Rasch model explicitly focuses on the measurement of construct validity. Each item is expected to make a meaningful contribution to the underlying construct or concept being investigated. Item fit statistics indicate the degree to which the observed response pattern for a particular item is consistent with the pattern predicted by the model, given the respondents' overall ability levels. Through the analysis of item fit, errors that occur during the calibration phase of instrument development can be clearly detected (Müller, 2020).

Instrument Reliability

The reliability of an instrument is evaluated to verify that it can consistently produce stable and dependable measurements. Within the Rasch modeling framework, reliability is assessed using three primary indices: person reliability, item reliability, and Cronbach's alpha (Sumintono & Widhiarso, 2015).

Tabel 6. Measurement of the Validity and Reliability of the Instrument

Aspect	Score
Pearson Measurement	
Person reliability	0.86
Separation	2,53
MNSQ	1.03
ZSTD	-0.75
Item Measurement	
Item reliability	0.97
separation	5.65
Cronbach alpha	0,89
Raw variance	45.7%
Unexplained variance 1st contrast	8.5%

As presented in Table 6, the item reliability index was 0.97, with an item separation of 5.65. This result indicates excellent item consistency and suggests that the items are sufficiently spread out along the measurement continuum to define a distinct hierarchy of difficulty. Reliability is expressed as a numerical score, typically ranging from 0 to +1.00, where a higher coefficient signifies greater reliability. High reliability is associated with a small degree of measurement error; consequently, the higher an instrument's reliability, the smaller its measurement error. Person reliability is an index that reflects the consistency of students' response patterns across the set of items. For the AI literacy instrument, the person reliability was found to be 0.86. This value indicates a good level of consistency in the students' response patterns.

The overall instrument reliability reported in Table 6 is 0.89, a value based on Cronbach's

alpha, which indicates a high degree of internal consistency. This strong value signifies a robust interaction between the students' abilities and the items. It implies that the students responded with a high degree of consistency and that the items were effective in measuring a wide range of proficiency levels. Therefore, the instrument can consistently measure the target construct and is expected to yield similar results upon repeated administration to the same subjects or to a sample with similar characteristics.

Unidimensionality

The assumption of item unidimensionality is tested to confirm that an instrument measures a single, intended construct, thereby providing evidence for its validity. Within the Rasch framework, this is evaluated using a Principal Component Analysis (PCA) of the residuals, which determines the proportion of variance explained by the primary construct (Ilhami & Hidayat, 2025). An instrument is considered unidimensional if it meets two conditions: first, the raw variance explained by the measures must be at least 20%, and second, the unexplained variance in the first contrast must not exceed 15% (Sumintono & Widhiarso, 2015). The results showed that the raw variance explained by the measures was 45.7%, indicating a strong primary dimension. Furthermore, the unexplained variance in the first contrast was 8.5%, well below the 15% threshold. These findings confirm that the instrument satisfies the criteria for unidimensionality, providing strong evidence that it effectively measures the intended construct.

The concept of unidimensionality within Rasch analysis posits that all items should relate to a single latent variable and measure it in a consistent direction. While this property is often taken for granted in quantitative social science research, it is crucial to empirically test this assumption. This is particularly important in novel contexts where the precise nature of the latent variable is still under investigation (Bond et al., 2021).

Rating Partial

The quality of an instrument's rating scale, which comprises the set of alternative response options, can be evaluated using rating scale diagnostics. The purpose of this analysis is to determine if the scale is functioning appropriately (able to discriminate between respondents of varying ability), reliable (consistent in this discrimination), and well-defined (respondents understand the distinctions between response categories) (Boone et al., 2014). The AI literacy instrument employed a six-point Likert-type scale with the following response categories: 6 (Strongly Agree), 5 (Agree), 4 (Somewhat Agree), 3 (Somewhat Disagree), 2 (Disagree), and 1 (Strongly Disagree). The diagnostic data for this rating scale analysis, generated from the partial credit model output, are presented in Table 7.

Table 7. Diagnostic Rating Test Results

Label	Category Label	Observed average	Andrich threshold
Strongly disagree	1	-0.12	None
Disagree	2	-0.49	-2.08

Somewhat disagree	3	-0.15	-6.0
Somewhat agree	4	0.51	-6.2
Agree	5	1.30	0.55
Strongly agree	6	1.76	2.75

For the response alternatives to be considered well-understood by respondents, both the observed average measures and the Andrich thresholds should increase monotonically with each successive rating scale category (Sumintono & Widhiarso, 2015). In this analysis, the observed average for respondents' answers showed a consistent increase. The specific observed average value for response option 1 was -0.12; for option 2, it was -0.49; for option 3, it was -0.15; for option 4, it was 0.51; for option 5, it was 1.30; and for option 6, it was 1.76. This result indicates that the AI literacy instrument for Islamic Religious Education (PAI) students functions appropriately for the sample and that the response categories were not confusing to the respondents.

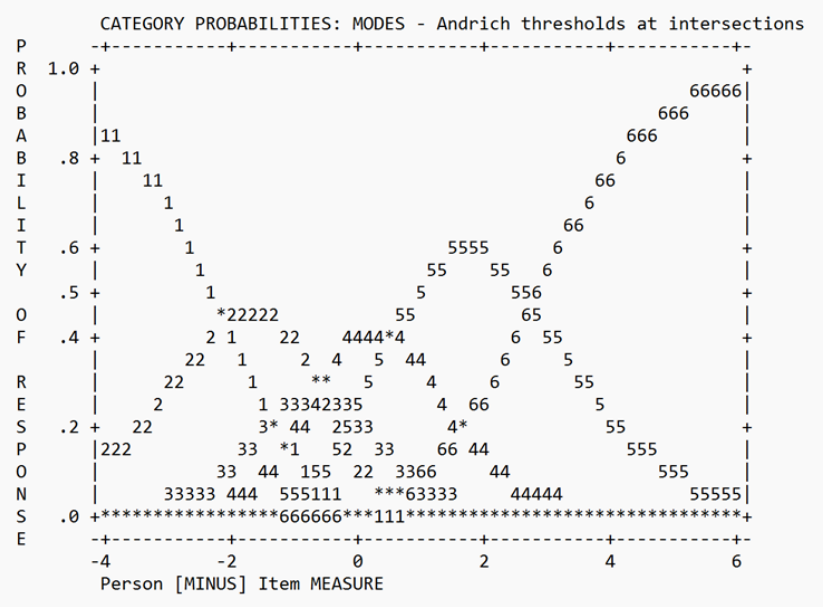

Figure 1. Probability Category Curve

Consecutive categories. These thresholds are important for understanding the extent to which differences between categories are acceptable and distinguishable to respondents. The Andrich threshold values monotonically progress from none toward negative logits and continue toward positive logits, indicating that the six response options provided are valid for the respondents (Sumintono & Widhiarso, 2015). Based on the analysis results, the threshold between the categories "Disagree" and "Slightly Disagree" was recorded at -2.08, while the threshold between "Slightly Disagree" and "Slightly Agree" was considerably larger, at -6.0. The substantial difference in thresholds between these categories indicates a significant gap in respondents' perceptions of the two categories. Similarly, the threshold between "Slightly Agree" and "Agree" was recorded at -6.2, also indicating a large difference. However, the threshold

between "Agree" and "Strongly Agree" was recorded at 0.55, suggesting that the difference between these categories is smaller, making the transition between them clearer for respondents. Overall, this measurement instrument demonstrates sufficiently distinct average measurements across each category (Figure 1). Nevertheless, the irregular threshold between the "Slightly Disagree" and "Slightly Agree" categories suggests that these categories may need to be revised to better differentiate respondents' levels of agreement.

Conclusion

Based on the results of the study, the AI literacy instrument meets the necessary assumptions for the Rasch Model, including fit, unidimensionality, rating scale, and reliability. The item reliability level was 0.97, while person reliability was 0.88, both of which fall into the good category. This instrument is reliable and demonstrates internal construct validity, as indicated by a raw variance explained measure of 45.7%. The AI literacy instrument can therefore be used to measure students' AI literacy. A total of 67 items were found to be valid, while 18 items should be considered for removal.

Acknowledgments

The research team would like to express their gratitude to the Research and Community Service Institute (LPPM) of UIN Sultan Syarif Kasim Riau and the Directorate of Islamic Higher Education (Diktis) of the Ministry of Religious Affairs of the Republic of Indonesia (Kemenag RI). This program was carried out through a research grant program for the 2025 fiscal year.

References

Abbas, T. (2023). Ethical Implications of AI in Modern Education: Balancing Innovation and Responsibility. *Social Sciences Spectrum*, *2*(1), 51–57.

Biagini, G., Cuomo, S., & Ranieri, M. (2023). Developing and Validating a Multidimensional AI Literacy Questionnaire: Operationalizing AI Literacy for Higher Education. In S. D., I. for E. T. National Research Council Via Ugo la Malfa 152, Palermo, T. D., I. for E. T. National Research Council Via Ugo la Malfa 152, Palermo, T. M., & I. I. Universita Roma Sapienza Automatica, e Gestionale "Antonio Ruberti", Via Ariosto, 25, Roma (Eds.), *CEUR Workshop Proceedings* (Vol. 3605). CEUR-WS.

Bond, T. G., Yan, Z., & Heene, M. (2021). Applying the Rasch Model:Fundamental Measurement in The Human Science. In *Applying the Rasch Model* (Fourth Edition). Routledge. https://doi.org/10.4324/9781410614575

Boone, W. J. (2016). Rasch Analysis For Instrument Development: Why, When, and How? *CBE—Life Sciences Education*, *15*(4), rm4.

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch Analysis In The Human Sciences. In *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Fadlelmula, F. K., & Qadhi, S. M. (2024). A systematic review of research on artificial intelligence in higher education: Practice, gaps, and future directions in the GCC. *Journal of University Teaching and Learning Practice*, *21*(6), 146–173.

Faisal Tariq Hasan, & Rafiq Rahman. (2024). Artificial Intelligence in Personalized Learning: A New Era of Education. *Proceeding of The International Conference of Inovation, Science, Technology, Education, Children, and Health*, *4*(2), 306–308. https://doi.org/10.62951/icistech.v4i2.127

Grassini, S. (2024). A Psychometric Validation of the PAILQ-6: Perceived Artificial Intelligence Literacy Questionnaire. *ACM International Conference Proceeding Series*.

https://doi.org/10.1145/3679318.3685359

Ilhami, A., & Hidayat, T. (2025). Psychometric validation of a research skills assessment for pre-service science teachers using the rasch model. *Thabiea: Journal Of Natural Science Teaching*, *8*(2), 149–159.

Kesuma, D. P., & Fransen, L. A. (2025). Implementasi Instrumen MAILS (Meta AI Literacy Scale) Untuk Pengukuran Tingkat Literasi AI Pada Mahasiswa Ilmu Komputer. *Jurnal Sistem Informasi, Manajemen dan Teknologi Informasi*, *3*(1), 25–36.

Metsämuuronen, J. (2023). Seeking the Real Item Difficulty: Bias-Corrected Item Difficulty and Some Consequences in Rasch and IRT Modeling. *Behaviormetrika*, *50*(1), 121–154.

Müller, M. (2020). Item fit statistics for Rasch analysis: can we trust them? *Journal of Statistical Distributions and Applications*, *7*(1), 5.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and Validation of the AI literacy Questionnaire: The affective, Behavioural, Cognitive And Ethical Approach. *British Journal of Educational Technology*, *55*(3), 1082–1104. https://doi.org/10.1111/bjet.13411

OECD. (2025). *Empowering Learners for the Age of AI* (Issue May). OECD Publishing.

Prasetya, W. A., & Pratama, A. T. (2023). Item Quality Analysis Using the Rasch Model to Measure Critical Thinking Ability in the Material of The Human Digestive System of Biology Subject in High School. *Jurnal Penelitian dan Evaluasi Pendidikan*, *27*(1), 76–91.

Rahayu, S. (2023). The Impact Of Artificial Intelligence on Education: Opportunities and Challenges. *Jurnal Educatio FKIP UNMA*, *9*(4), 2132–2140.

Sari, H. P., Ilhami, A., & Rasool, S. (2025). Implementation of Project Based Learning Based Local Wisdom to Enhance Students' Critical Thinking in Higher Education. *Journal of Contemporary Islamic Education*, *5*(1), 15–27.

Shi, L. (2025). Assessing Teachers' Generative Artificial Intelligence Competencies: Instrument Development And Validation. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13684-5

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan* (1st ed.). Trim Komunikata Publishing House.

UNESCO. (2024). *AI Competency Framework for Students*. United Nations Educational Scientific and Cultural Organization. https://doi.org/10.54675/jkjb9835

Wilton, L., Ip, S., Sharma, M., & Fan, F. (2022). Where is the AI? AI Literacy for Educators. *International Conference on Artificial Intelligence in Education*, 180–188.

Zhang, J., & Zhang, Z. (2024). AI in Teacher Education: Unlocking New Dimensions in Teaching Support, Inclusive Learning, and Digital Literacy. *Journal of Computer Assisted Learning*, *40*(4), 1871–1885.

Zootzky, G., & Pfeiffer, A. (2024). Educational transformation through AI: Preparing for a new era of learning. *INTED2024 Proceedings*, 1202–1207.