Developing a Guttman Scale to Measure Teachers' Attitudes Toward the Use of Artificial Intelligence in Education Based on Item Response Theory

تطوير مقياس جتمان لاتجاهات المعلمين نحو توظيف الذكاء الاصطناعي في التعليم وفقاً لنظرية الاستجابة للفقرة

10.35781/1637-000-204-007

Sharif Alsoudi[1], Mohammed Al Tobi[2], Abir Al-Nassri[3]

1Department of Psychology, Arts & Humanities College, A'Sharqiyah University, Ibra, Oman.
Email: sharif.alsoudi@asu.edu.om
2Department of Education, Arts & Humanities College, A'Sharqiyah University, Ibra, Oman.
Email: altobi_mohammed@yahoo.com
3Ministry of Education, Oman.
Email: 2319106@asu.edu.om
Corresponding author: Sharif Alsoudi

Abstract:

This study developed and validated a Guttman-scale-based Item Response Theory (IRT) model to measure teachers' attitudes toward using Artificial Intelligence (AI) in education. The scale was administered to 523 teachers in Oman, selected through a diverse sampling method across eight educational regions. Using Rasch model analysis and exploratory factor analysis, the 11-item scale demonstrated strong psychometric properties, including unidimensionality, high reliability (Cronbach's α = .846; Rasch person reliability = .98), and excellent model fit (CFI = .981, RMSEA = .044). The scale revealed a clear hierarchical structure of attitudes, progressing from basic acceptance of AI concepts to transformative integration in teaching practices. Results indicated that the scale effectively measures attitude intensity along a continuum, with items covering a wide range of difficulty parameters (-2.29 to 2.53 logits). The scale provides educational researchers and practitioners with a validated tool to assess teacher readiness for AI integration, design targeted professional development programs, and inform policy decisions on AI implementation in educational contexts.

Keywords: Artificial Intelligence, Teacher Attitudes, Guttman Scale, Rasch Model, Psychometric Properties, Scale Development.

ISSN : 2410-1818

تطوير مقياس جتمان لاتجاهات المعلمين نحو توظيف الذكاء الاصطناعي في التعليم وفقاً لنظرية الاستجابة للفقرة

شريف السعودي[1]، محمد التوبي[2]، عبير الناصرية[3]

[1]قسم علم النفس، كلية الآداب والعلوم الإنسانية، جامعة الشرقية، إبراء، سلطنة عُمان.
البريد الإلكتروني: sharif.alsoudi@asu.edu.om
[2]قسم التربية، كلية الآداب والعلوم الإنسانية، جامعة الشرقية، إبراء، سلطنة عُمان.
البريد الإلكتروني: altobi_mohammed@yahoo.com
[3]وزارة التربية والتعليم، سلطنة عُمان.
البريد الإلكتروني: 2319106@asu.edu.om

الملخص

هدفت الدراسة إلى تطوير مقياس وفق طريقة جتمان ونظرية الاستجابة للفقرة (IRT)، والتحقق من خصائصه السيكومترية لقياس اتجاهات المعلمين نحو استخدام الذكاء الاصطناعي في التعليم. طُبّق المقياس على عينة مكونة من (523) معلماً ومعلمة في سلطنة عُمان، اختيروا باستخدام أسلوب معاينة يراعي التنوع عبر ثماني مناطق تعليمية. وباستخدام نموذج راش والتحليل العاملي الاستكشافي، أظهرت نتائج تحليل المقياس، المكون من (11) فقرة، تمتعه بخصائص سيكومترية جيدة، تمثلت في أحادية البعد، وارتفاع مؤشرات الثبات، إذ بلغ معامل ألفا كرونباخ (0.846)، ومعامل ثبات الأفراد وفق نموذج راش (0.98)، فضلًا عن مؤشرات مطابقة جيدة للنموذج (CFI = 0.981, RMSEA = 0.044). وكشفت النتائج عن وجود بنية هرمية واضحة للاتجاهات، تتدرج من التقبل الأساسي لمفاهيم الذكاء الاصطناعي إلى دمجه بصورة تحولية في الممارسات التدريسية. كما أظهرت النتائج فاعلية المقياس في قياس شدة الاتجاهات على متصل متدرج، إذ غطّت فقراته مدى واسعًا من معاملات الصعوبة تراوح بين (–2.29 و2.53) لوجيت. وتوفر الدراسة أداة قياس تتمتع بمؤشرات صدق وثبات مناسبة، يمكن للباحثين والممارسين التربويين الاستفادة منها في تقييم استعداد المعلمين لدمج الذكاء الاصطناعي في التعليم، وتصميم برامج التنمية المهنية الموجهة، ودعم القرارات والسياسات التربوية المتعلقة بتوظيف الذكاء الاصطناعي في البيئات التعليمية.

الكلمات المفتاحية: الذكاء الاصطناعي، اتجاهات المعلمين، مقياس جتمان، نموذج راش، الخصائص السيكومترية، تطوير المقاييس.

ISSN : 2410-1818

Introduction

The rapid integration of Artificial Intelligence (AI) into education represents a paradigm shift, promising to redefine traditional pedagogical approaches. AI applications, such as intelligent tutoring systems, adaptive learning platforms, and automated assessment tools, are no longer futuristic concepts but present-day realities with the potential to personalize learning, streamline administrative tasks, and foster critical thinking skills (Afif, 2023; Ghamdi, 2024; UNESCO, 2023). However, the ultimate efficacy of these technological advancements is not solely determined by their algorithmic sophistication. Still, it is profoundly mediated by a critical human factor: the teachers who must integrate them into their practice.

Teachers' attitudes—encompassing their perceptions, acceptance, and readiness—are a pivotal determinant in the successful adoption and sustainable implementation of AI in education. A positive disposition toward AI correlates strongly with a teacher's willingness to engage in professional development, experiment with innovative teaching strategies, and leverage AI to enhance student feedback and learning outcomes (Al-Muslim, 2023; Çayak, 2024). Conversely, apprehension stemming from ethical concerns about data privacy and algorithmic bias, fears of deprofessionalization, or simply a lack of technological self-efficacy can create significant barriers to adoption, leading to resistance and ineffective use (Ibrahim, 2024; Al-Dulaimi & Al-Amiri, 2024). Therefore, moving beyond anecdotal evidence to a precise, scientifically rigorous measurement of these attitudes is imperative for educational stakeholders to design targeted interventions, effective training programs, and supportive policies.

The field of psychometrics offers robust methodologies for such measurement. The Guttman scale, a classic approach, provides a unique advantage by conceptualizing attitudes as a cumulative, hierarchical structure. In the context of AI acceptance, this could translate to a scale where agreement with a complex item (e.g., "I am comfortable allowing an AI system to generate personalized learning paths for my students autonomously") logically implies agreement with all preceding, less intense items (e.g., "I believe AI is useful for grading multiple-choice tests"). This allows researchers to not only measure the direction of an attitude (positive/negative) but also its intensity and the specific point at which acceptance turns into resistance.

While the Guttman scale provides a powerful structural model, modern measurement theory enhances it with statistical rigor. Item Response Theory (IRT), particularly the Rasch model, allows for the transformation of ordinal survey responses into interval-level measurements. This enables a more precise calibration of "item difficulty" (in this case, the level of AI acceptance an item represents) and "person ability" (a teacher's level of acceptance). Rasch analysis provides valuable diagnostics, identifying misfitting items that do not conform to the model's expectations and flagging potential biases through Differential Item Functioning (DIF), thereby significantly improving the instrument's validity and reliability (Bond & Fox, 2015; Wang, 2024). The combined use of Guttman scaling and Rasch

ISSN : 2410-1818

modeling has proven effective in constructing robust attitudinal measures in Arab educational contexts, as shown in recent studies (e.g., Alsoudi & AlHarthy, 2024; Alsoudi & Abu Shindi, 2023).

The synergy of Guttman's conceptual hierarchy and Rasch's probabilistic modeling creates a comprehensive framework for scale development. This integrated approach is not only methodologically sound but has also been successfully validated for measuring complex constructs in the region, providing a strong precedent for the current study (Alsoudi & AlHarthy, 2024). This dual approach is especially crucial for investigating teachers' attitudes toward AI in Middle Eastern educational systems, where cultural, institutional, and resource-related factors may shape perceptions in ways not fully captured by generic surveys (Al-Shammari, 2024). Currently, a significant gap exists in the literature: while advanced psychometrics has been applied to other domains, such as critical thinking (Alsoudi & Abu Shindi, 2023), its application to AI attitudes in education remains limited. Many relevant studies still rely on simple descriptive surveys or Likert-type scales that lack the diagnostic power to reveal the underlying, cumulative structure of attitudes.

To address this gap, this study proposes developing and validating a Guttman scale, grounded in Item Response Theory, to measure teachers' attitudes toward using AI in education. The resulting instrument will provide a nuanced, reliable, and valid tool that can diagnose not only whether teachers are receptive to AI, but also the extent of their support and the specific stages at which their support wavers. The findings are expected to yield actionable insights for teacher professional development, curriculum design for AI integration, and policy formulation, ultimately facilitating a more effective and human-centric adoption of artificial intelligence in classrooms.

Method

Participants

The study employed a diverse sample of 523 in-service teachers from eight educational regions across Oman. Data collection proceeded with official approval from the Omani Ministry of Education, which facilitated voluntary recruitment via an online survey. Prior to participation, all respondents provided informed consent electronically. The sample was designed to capture a broad spectrum of the teaching workforce, encompassing variations in gender, academic qualifications, years of professional experience, subject specialization, and the educational stage at which they teach.

As detailed in Table 1, the sample composition reflects key characteristics of the teacher population. Notably, most participants held a bachelor's degree and had substantial teaching experience, with over 60% reporting ten or more years in the profession. The sample also included a balanced representation of both scientific and humanities specializations, as well as teachers from all three core educational cycles. This strategic diversity, supported by the Ministry, enhances the sample's

ISSN : 2410-1818

representativeness and strengthens the generalizability of the study's findings to the broader context of Omani teachers.

Table 1. Demographic Characteristics of the Teachers Sample (N = 523)

Variable	Category	N	%
Sex	Male	225	43.02
	Female	298	56.98
Qualification	Diploma	76	14.53
	Bachelor's Degree	379	72.47
	Postgraduate	68	13.00
Years of Experience	Less than 5	98	18.7
	5–9	94	18.0
	10–15	130	24.9
	More than 15	201	38.4
Specialization	Humanities	212	40.5
	Scientific	311	59.5
Educational Stage	Cycle 1 (Grades 1-4)	190	36.3
	Cycle 2 (Grades 5-9)	226	43.2
	Cycle 3 (Grades 10-12)	107	20.5

Instrument Development and Procedures

The development of the Guttman scale was initiated with a comprehensive literature review, forming the basis for an initial 15-item pool. These items were meticulously designed to encompass the cognitive, affective, and behavioral dimensions of attitudes, structured to represent a hierarchical continuum from the most basic level of favorability towards AI use, such as recognizing its utility for routine tasks, to the most advanced level, such as endorsing its role in transformative pedagogical practices. To rigorously establish content validity and item clarity, the preliminary scale underwent a formal evaluation by a panel of eight experts, including six specialists in psychometrics and two in educational technology. The experts assessed each item's relevance, clarity, and appropriateness. This process led to the removal of two items (A2, A9); one was deleted due to consensus on its lack of applicability in the current context, specifically regarding AI's role in diagnosing learning disabilities, and another was removed for substantive redundancy with other items addressing teacher performance improvement. The outcome was a refined and validated 13-item instrument.

Subsequently, a pilot study was administered to a sample of 25 teachers to verify the instrument's linguistic clarity and practical comprehensibility. Feedback was unequivocally positive: all participants confirmed that all 13 items were clear and understandable, and no ambiguities or response difficulties were reported. This successful pilot confirmed the instrument's readiness for widespread administration.

ISSN : 2410-1818

After these development and validation stages, the Ministry of Education granted official approval. The final 13-item scale was then deployed electronically to the main study sample of 523 teachers across the eight educational regions in Oman.

Data Analysis

The data analysis was conducted using a two-stage analytical approach to validate the psychometric properties of the developed Guttman scale, employing both classical and item response theory frameworks.

Initial data screening and classical psychometric analyses were performed using IBM SPSS Statistics (Version 30). An Exploratory Factor Analysis (EFA) utilizing the Principal Axis Factoring method was conducted to investigate the fundamental unidimensionality of the scale. Furthermore, the instrument's internal consistency reliability was estimated using Cronbach's Alpha coefficient.

The primary analysis focused on evaluating the scale's fit to the Rasch model using WINSTEPS software (Version 5.9.1). This comprehensive assessment included verifying the key assumptions of the model: unidimensionality was examined through raw residual variance in eigenvalue units, the requirement of equal item discrimination was tested using the likelihood-ratio test for parallel ICCs/IRFs, and local independence was assessed by analyzing correlations between standardized residuals. Model fit was evaluated at the item and person level using infit and outfit mean-square statistics (MNSQ) and standardized z-values (ZSTD). The analysis also calibrated item difficulty parameters, examined point-measure correlations, and calculated person and item separation indices and reliability coefficients to establish the scale's precision and discriminative power.

Finally, the overall model-data fit for the unidimensional Rasch model was assessed using the R programming environment with the haven and TAM packages. The analysis evaluated model fit using standard indices: Chi-square to degrees of freedom (χ^2/df), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Additionally, the scalability of the Guttman structure was examined through Loevinger's scalability coefficients (H) along with violation analysis. The coefficients of reproducibility and scalability were also calculated to further validate the cumulative nature of the attitude scale.

Results

Exploratory Factor Analysis and Unidimensionality

An exploratory factor analysis (EFA) using principal axis factoring was conducted on the final 11-item scale to verify the fundamental unidimensionality assumption of the Rasch model. The determinant value of .034 confirmed the absence of multicollinearity issues in the data. The Kaiser-Meyer-Olkin (KMO) measure verified high sampling adequacy with a value of .882, and Bartlett's test of

ISSN : 2410-1818

sphericity was statistically significant ($\chi^2 = 1752.730$, df = 55, $p < .001$), confirming the data's suitability for factor analysis.

The EFA results conducted on the final 11-item version (after removing two items from the original 13-item pool) supported the essential unidimensionality assumption for Rasch analysis, with a single dominant factor emerging (eigenvalue = 4.517) that accounted for 35.38% of the total variance. All retained items demonstrated statistically significant factor loadings exceeding 0.50 (range: 0.519-0.724), confirming their strong relationship with the underlying construct. The scree plot (Figure 1) further validated the unidimensional structure by showing a clear break between the first factor (eigenvalue = 4.517) and subsequent factors, with the characteristic "elbow" shape indicating a single dominant dimension. The removal of two items - "I get excited about using AI in teaching" (A6) and "I believe AI can provide personalized education for each student" (A13) - during preliminary analysis due to insufficient factor loadings (0.28 and 0.19, respectively) substantially enhanced both the scale's unidimensional structure and its internal consistency, resulting in a more coherent measurement instrument.

Table 2. Factor Loadings and Variance Explained

Item	Factor Loading
A1	0.704
A3	0.720
A4	0.553
A5	0.724
A7	0.610
A8	0.682
A10	0.617
A11	0.613
A12	0.571
A14	0.696
A15	0.519
Initial Eigenvalue	4.517
% of Variance	35.38%

ISSN : 2410-1818

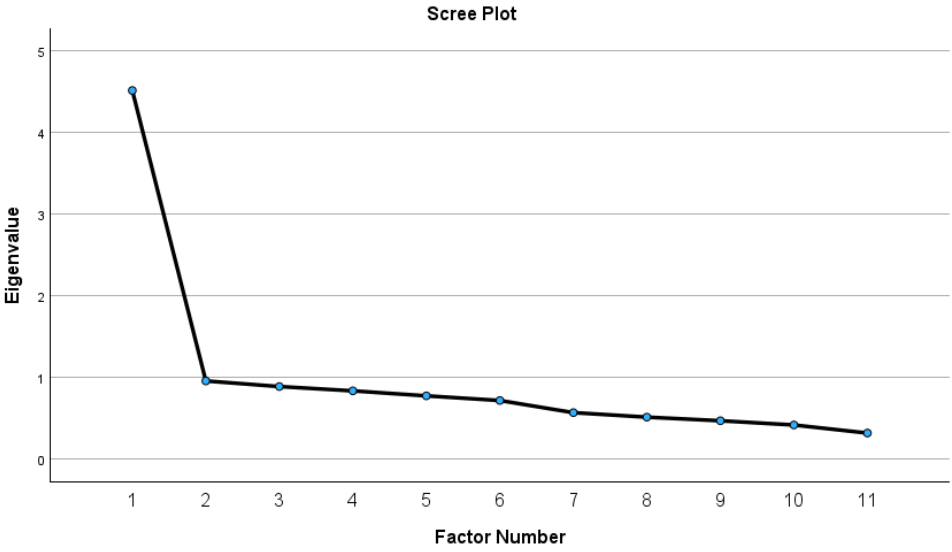


Figure 1. Scree Plot of Eigenvalues

Further evidence for unidimensionality was established through Rasch residual analysis. As shown in Table 3, the essential unidimensionality index reached 67.60%, significantly exceeding the minimum acceptable threshold of 50% recommended for Rasch measurement (Linacre, 2024). The Rasch measurement model explained 42.56% of the total variance, with items accounting for 23.06% and persons explaining 19.50%. Most importantly, the variance explained by the first contrast in the residuals (9.40%) was substantially lower than the 15% criterion suggested by Reckase (1979) and confirmed by Raîche (2005) for identifying secondary dimensions. All subsequent contrasts showed progressively smaller values, with none approaching the critical eigenvalue size of 3 (Smith et al., 1994). This pattern confirms that no substantial secondary dimension exists in the data, thereby strongly supporting the scale's unidimensional structure.

Table 3. Standardized Residual Variance Explained

Variance Type	Value	Percentage
Total Variance in Observations	17.39	100.00%
Explained Variance	7.40	42.56%
Explained by Measures	4.01	23.06%
Explained by Persons	3.39	19.50%
Unexplained Variance (Total)	9.99	57.44%
Unexplained in 1st Contrast	1.64	9.40%

ISSN : 2410-1818

Variance Type	Value	Percentage
Unexplained in 2nd Contrast	1.43	8.20%
Unexplained in 3rd Contrast	1.22	7.00%
Unexplained in 4th Contrast	1.20	6.90%
Unexplained in 5th Contrast	1.14	6.60%
Essential Unidimensionality (Rasch/Shared Variance)	-	67.60%

Rasch Model Fit Analysis

The Rasch model demonstrated excellent fit to the data across all goodness-of-fit indices (Table 4). The non-significant chi-square statistic relative to degrees of freedom ($\chi^2/df = 1.97$) fell well below the recommended threshold of 3 (Kline, 2023). Both incremental fit indices exceeded the 0.95 excellence criterion (TLI = 0.977; CFI = 0.981), indicating superior model fit compared to a null model (Hu & Bentler, 1999). The absolute fit indices also showed outstanding results, with RMSEA = 0.044 (90% CI [0.030, 0.058]) below the 0.06 cutoff for close fit (Browne & Cudeck, 1993), and SRMR = 0.039 meeting the strict criterion of ≤ 0.05 (Hu & Bentler, 1999). This comprehensive pattern of fit indices provides strong evidence that the unidimensional Rasch model adequately represents the underlying structure of the attitude scale.

Table 4. Goodness-of-Fit Indices for the Rasch Model

Fit Index	Value	Acceptable Range
Chi-square (df)	86.83 (44)	-
Chi-square/df ratio (χ^2/df)	1.97	$\chi^2/df \leq 3$
Tucker-Lewis Index (TLI)	0.977	$0.90 \leq TLI < 1.00$
Comparative Fit Index (CFI)	0.981	$0.90 \leq CFI < 1.00$
RMSEA	0.044	$0.00 \leq RMSEA \leq 0.08$
SRMR	0.039	$0.00 \leq SRMR \leq 0.05$

Item Fit to the Rasch Model

The item fit statistics presented in Table 5 demonstrate excellent alignment with the Rasch model expectations. All items exhibited acceptable infit and outfit mean-square (MNSQ) values within the recommended range of 0.5 to 1.5 (Bond & Fox, 2015), with mean values of 0.98 (infit) and 1.00 (outfit). The standardized fit statistics (ZSTD) for all items fell within the acceptable range of -2.0 to +2.0 (Linacre, 2020), indicating no significant misfit. The point-measure correlations ranged from 0.42 to 0.66 (mean = 0.54), all exceeding the minimum threshold of 0.40, confirming that each item contributes positively to the measurement of the underlying construct (Wright & Linacre, 1989). These results provide strong

ISSN : 2410-1818

evidence for the psychometric quality of all 11 items and their appropriate functioning within the Rasch measurement framework.

Table 5. Item Fit Statistics for the Rasch Model

Item	SE	INFIT		OUTFIT		Correlation
		MNSQ	ZSTD	MNSQ	ZSTD	
A1	0.30	0.85	-0.55	0.87	-0.09	0.59
A3	0.19	0.93	-1.02	0.83	-0.73	0.66
A4	0.12	1.10	1.33	1.41	1.57	0.60
A5	0.16	1.00	0.04	1.04	0.32	0.50
A7	0.23	1.11	0.70	1.07	0.31	0.53
A8	0.15	0.83	-0.79	0.80	-1.24	0.47
A10	0.15	0.83	-0.84	0.88	-1.04	0.52
A11	0.19	0.95	-0.33	0.84	-0.90	0.42
A12	0.12	0.95	-0.69	0.85	-1.16	0.55
A14	0.18	1.04	0.34	1.01	0.14	0.56
A15	0.12	1.21	1.79	1.40	1.94	0.58
Mean	0.17	0.98	0.00	1.00	-0.08	0.54
SD	0.05	0.12	0.90	0.21	1.03	0.06

Local Independence

Local independence was assessed by examining the correlations between standardized residuals. The analysis revealed that all residual correlations fell within an acceptable range of -0.20 to +0.14, with none exceeding the absolute value of 0.30 recommended by Christensen et al. (2017). The highest correlation observed was 0.14 between items A3 and A7, followed by -0.20 between A11 and A15, both well below the critical threshold. This pattern of low residual correlations confirms that the local independence assumption of the Rasch model is satisfied, indicating that items function independently after accounting for the dominant latent trait (Wang et al., 2018). The absence of substantial local dependence supports the appropriateness of the Rasch model for this dataset.

Equal Item Discrimination Test

The assumption of equal item discrimination was tested using the likelihood-ratio test for parallel item characteristic curves. The results showed a non-significant chi-square value ($\chi^2 = 10.372$, $df = 10$, $p = 0.408$), indicating that the hypothesis of equal discrimination parameters across all items cannot be rejected. The item discrimination parameters ranged from 0.79 to 1.25 with a mean of 1.03 (SD = 0.12), demonstrating reasonably uniform discrimination power. These findings satisfy the

ISSN : 2410-1818

fundamental Rasch model requirement of equal item discrimination, confirming that all items contribute equally to measuring the underlying construct and supporting the appropriateness of the Rasch model over more complex item response theory models that allow for varying discrimination parameters (Bond et al., 2021).

Scalability and Guttman Scale Properties

The scalability analysis presented in Table 6 demonstrates strong evidence for the Guttman scale structure. The coefficient of reproducibility (0.971) substantially exceeds the minimum criterion of 0.90, indicating excellent predictability of response patterns (Müller, 2023). Similarly, the coefficient of scalability (0.698) surpasses the recommended threshold of 0.60, confirming the cumulative hierarchy of items (Loevinger, 1948). Item scalability coefficients (H) ranged from 0.40 to 0.58 (mean = 0.49), all meeting the acceptable minimum of 0.30. The low violation rate (mean = 4%) and minimal Z-statistics further support the scale's cumulative nature. These results collectively validate the instrument's Guttman scaling properties, demonstrating that items form a reliable hierarchical continuum measuring teachers' attitudes toward AI in education.

Table 6. Scalability Coefficients and Guttman Scale Indices

Item	Scalability Coefficient (H)	Number of Comparisons	Number of Violations	Violation Percentage	Maximum Z
A1	0.58	37	0	0%	0.00
A3	0.54	37	1	3%	1.81
A4	0.49	30	2	7%	1.33
A5	0.44	36	2	6%	1.68
A7	0.45	37	0	0%	0.00
A8	0.55	35	2	6%	1.68
A10	0.52	36	2	6%	1.25
A11	0.40	37	1	3%	0.51
A12	0.53	30	2	7%	1.33
A14	0.41	37	2	5%	1.81
A15	0.48	30	0	0%	0.00
Mean	0.49	34.73	1.27	4%	1.04
Coefficient of Reproducibility				0.971	
Coefficient of Scalability				0.698	

Reliability and Separation

Reliability and separation were assessed using multiple methodologies. Rasch analysis revealed an item reliability coefficient of 0.98, indicating exceptional stability of the item structure across potential samples (Fisher, 1992). The separation

ISSN : 2410-1818

analysis showed a high value of 7.20, demonstrating the scale's ability to distinguish more than seven distinct levels of attitudes toward AI (Wright & Masters, 1982). Additionally, Cronbach's alpha coefficient reached 0.846, exceeding the minimum acceptable threshold (0.70) and confirming the scale's internal consistency (Tavakol & Dennick, 2011). These collective results indicate that the scale possesses strong psychometric properties, including both reliability and sensitivity in measuring attitudes.

Item Hierarchy and Scale Cumulatively

Table 7 presents the final hierarchical structure of the Guttman scale, demonstrating a clear cumulative pattern of attitudes toward AI in education. The item difficulty parameters span a wide range from -2.29 to 2.53 logits, creating a comprehensive measurement continuum. The hierarchical ordering reveals a logical progression from basic acceptance of AI concepts (easiest items: A1, A7) to transformative applications (most difficult items: A15, A4). This perfect difficulty gradient confirms the scale's cumulative nature, where endorsement of higher-level items implies endorsement of all lower-level items, fulfilling the core requirement of Guttman scaling (Guttman, 1950). The smooth transition from supportive attitudes to practical implementation and, finally, to role transformation validates the theoretical construct and provides strong evidence of the scale's construct validity (Bond & Fox, 2015).

Table 7. Item Hierarchy, Difficulty Parameters, and Endorsement Rates

Rank	Item	Item Content	N Endorsed	% Endorsed	Difficulty
1	A15	I believe AI can replace teachers in some teaching tasks.	304	58.13%	2.53
2	A4	I support using AI for grading assignments and tests.	362	69.22%	1.66
3	A12	I see the need for AI in continuous student assessment.	372	71.13%	1.51
4	A10	I support integrating AI programs into core curricula.	411	78.59%	0.66
5	A8	I believe AI helps reduce teachers' administrative burdens.	452	86.42%	0.03
6	A5	I believe AI can reduce teachers' instructional workloads.	453	86.62%	0.01

ISSN : 2410-1818

Rank	Item	Item Content	N Endorsed	% Endorsed	Difficulty
7	A14	I believe teachers should redesign their traditional roles to benefit from AI.	476	91.01%	-0.65
8	A3	I trust that AI can help students achieve better learning outcomes.	477	91.20%	-0.68
9	A11	I believe in using AI as a tool for preparing lesson plans.	481	91.97%	-0.83
10	A7	I believe AI can make learning more enjoyable.	493	94.26%	-1.35
11	A1	I support including AI concepts in education.	507	96.94%	-2.29

Rasch Model Graphical Analyses

Figure 2 displays the Wright Map, illustrating the joint distribution of persons and items along the latent variable continuum. The map demonstrates excellent targeting, with item difficulties well-matched to person abilities across the measurement range (approximately -3 to +4 logits). The clear hierarchical ordering of items confirms the scale's cumulative structure, with A15 "AI can replace teachers" representing the most difficult endorsement and A1 "support learning about AI concepts" being the easiest. The concentration of persons around the mean (0 logits) with adequate spread across the continuum indicates the scale effectively captures the attitudinal variation in the teacher population.

Figure 3 shows the Test Information Function, revealing optimal measurement precision across a broad range of the latent trait. The peak information (approximately 4.0) around zero logits indicates superior precision near the mean attitude level, while maintaining robust precision (information > 2.0) across the -2 to +2 logit range. This pattern confirms the scale's effectiveness in measuring teachers with moderate attitude levels while providing adequate precision for those with more extreme positions, supporting its utility for diverse educational applications.

Figure 4 presents the Category Probability Curves for a typical dichotomous item, demonstrating excellent response category functioning. The clear threshold at approximately zero logits, combined with the steep transition between categories, indicates strong item discrimination and proper functioning of the response format. The monotonically increasing "Agree" curve and corresponding decreasing "Disagree" curve validate the logical response pattern, where higher attitude levels predict higher endorsement probabilities.

Collectively, these graphical analyses provide strong visual confirmation of the scale's adherence to Rasch model requirements, its robust measurement

ISSN : 2410-1818

properties, and its effectiveness as a unidimensional measure of teachers' attitudes toward AI in education.

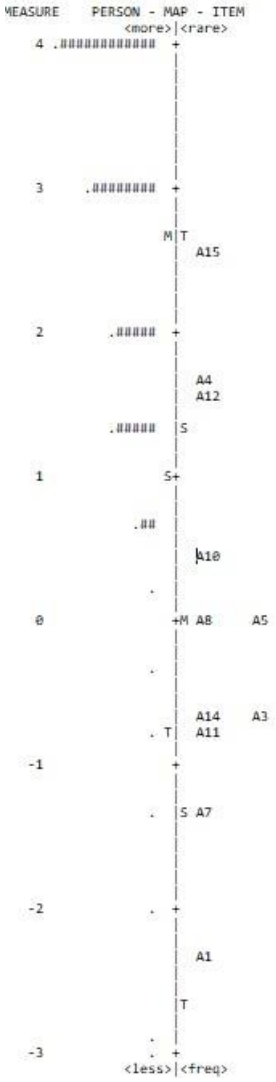


Figure 2. Wright Map (Person-Item Distribution)

ISSN : 2410-1818

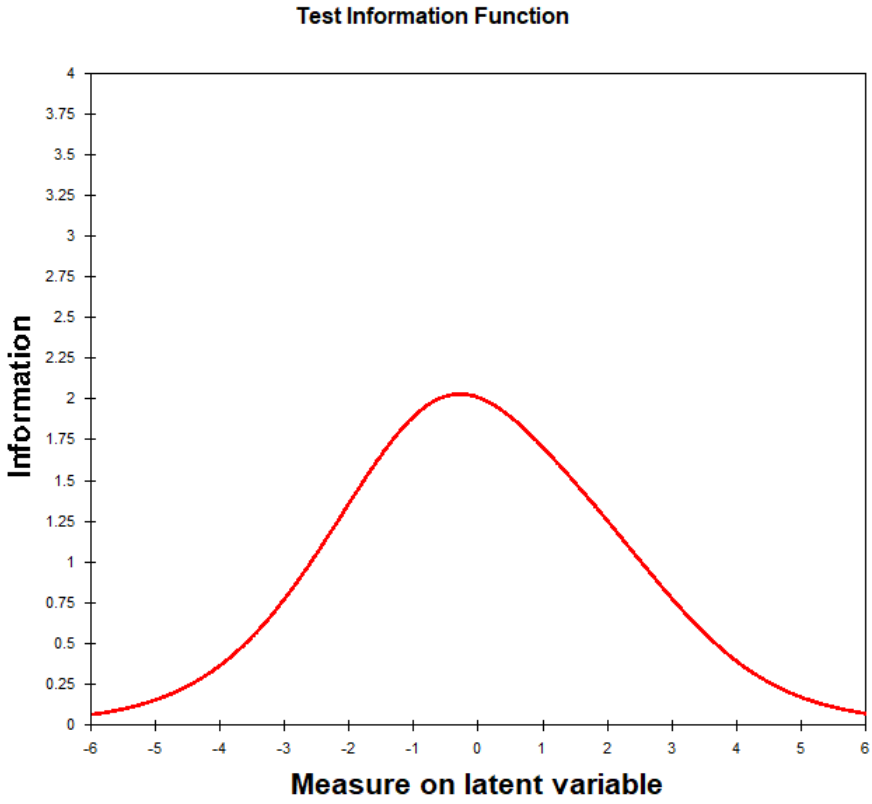


Figure 3. Test Information Function

ISSN : 2410-1818

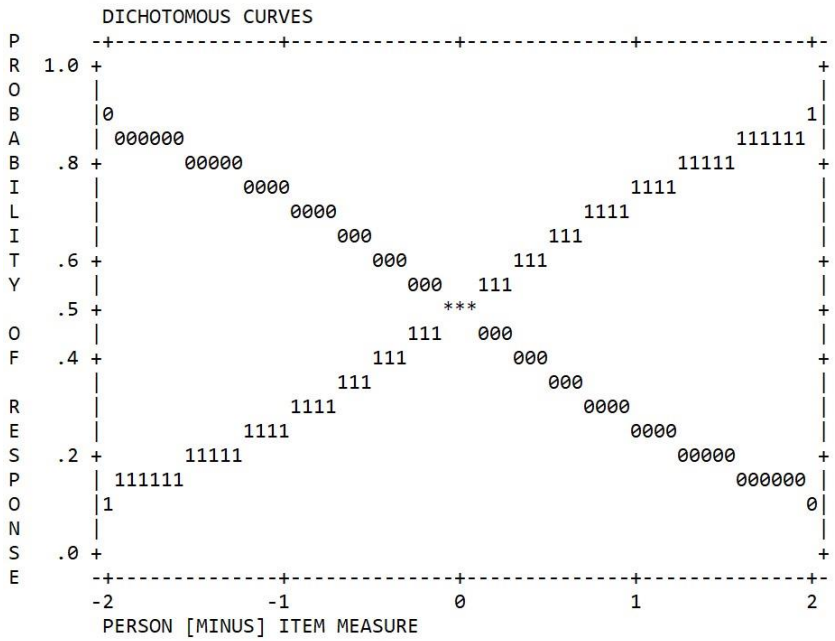


Figure 4. Category Probability Curves for a Dichotomous Item

ISSN : 2410-1818

Discussion

The present study successfully developed and validated a unidimensional Guttman scale to measure teachers' attitudes toward the use of artificial intelligence in education, based on the Rasch model. The findings demonstrated strong evidence for the scale's psychometric properties, including excellent model-data fit, high reliability, and a clear hierarchical structure. The scale exhibited robust unidimensionality, with both exploratory factor analysis and Rasch residual analysis confirming that a single dominant construct underlies the items. The item hierarchy revealed a logical progression from basic conceptual acceptance of AI to its transformative role in teaching, ultimately reaching the most challenging endorsement of AI as a potential partial replacement for teachers. This cumulative structure aligns with the core principle of Guttman scaling, confirming that the scale effectively measures the intensity of teachers' attitudes along a well-defined continuum.

The clear hierarchical structure emerging from the scale calibration provides valuable insights into teachers' psychological acceptance of AI. The finding that acceptance of AI's role in automating administrative tasks (e.g., A5, A8) precedes acceptance of its pedagogical role (e.g., A3, A7), which in turn precedes the endorsement of its transformative potential (A14) and finally its role in partial replacement (A15), aligns with established technology adoption models. This progression mirrors the Technology Acceptance Model (TAM) framework, where perceived usefulness for simple tasks typically precedes acceptance for core professional functions (Davis, 1989; Scherer et al., 2019). The most difficult item (A15), touching on role replacement, resonates with literature on "technostress" and professional identity threat among educators, where technologies perceived as challenging core professional autonomy face the strongest resistance (Jisc, 2021; Wang et al., 2023).

The scale's strong psychometric properties, particularly its high-reliability indices (Cronbach's Alpha = 0.846; Rasch Person Reliability = 0.98) and excellent model-fit (CFI = 0.981, RMSEA = 0.044), place it among the most robust instruments in the field of educational technology attitude measurement. These results surpass those reported in earlier studies on teacher technology acceptance (e.g., Teo, 2019; Hsu & Lin, 2023) and align with the increasing methodological rigor called for by recent systematic reviews (Bond et al., 2020). The wide measurement range and high precision (as shown by the Test Information Function) indicate that the scale is particularly adept at detecting subtle differences in attitude, making it a sensitive tool for evaluating professional development interventions. This finding is crucial, as it addresses the limitation of "ceiling effects" noted in previous attitudinal measures where most respondents cluster at the positive end of the scale (Zheng et al., 2022).

The compelling evidence for unidimensionality from both EFA and Rasch residual analysis (Raw Variance explained by measures = 23.06%; Unexplained

ISSN : 2410-1818

variance in 1st contrast = 9.4%) strongly suggests that teachers' attitudes toward AI in education form a coherent, single construct. This finding challenges the multidimensional frameworks sometimes proposed for technology acceptance, which separate constructs like perceived usefulness, ease of use, and behavioral intention (Venkatesh & Davis, 2000; Granić & Marangunić, 2023). Instead, it supports a more parsimonious view where these facets are interconnected manifestations of a core attitudinal disposition, at least in the specific context of AI. This aligns with a growing body of evidence suggesting that for transformative technologies like AI, attitudes can crystallize into a general, overarching evaluation that guides specific perceptions (Zhang et al., 2023).

Conclusion and Implications

This study successfully developed and validated a rigorous Guttman scale to measure teachers' attitudes toward AI in education using Rasch measurement theory. The 11-item scale demonstrates strong psychometric properties, unidimensionality, and a clear hierarchical structure that reveals the cumulative nature of AI acceptance—from basic conceptual understanding to transformative integration. The scale provides a reliable and valid tool for capturing the progression of teacher attitudes toward AI.

Practical implications emerge directly from these findings. For teachers and school leaders, the scale can serve as a self-assessment tool to identify personal readiness levels and target specific areas for professional growth. For policymakers, it offers an evidence-based framework for designing differentiated training programs that address teachers' varying acceptance levels. For teacher training institutions, the scale provides a diagnostic tool to customize curriculum development for AI integration in education. By pinpointing specific transition points in the acceptance continuum, the scale enables targeted interventions that can effectively support teachers through the challenging process of AI adoption in their professional practice.

Limitations and Future Research

Despite the robust psychometric properties demonstrated in this study, several limitations warrant acknowledgment. The sample, though diverse in teaching backgrounds, was drawn exclusively from one country (Oman), which may limit the generalizability of the findings to other cultural and educational contexts. Convenience sampling, while administratively practical, is another limitation. Furthermore, the cross-sectional design provides a snapshot of attitudes. It cannot capture how attitudes evolve as AI technologies continue to develop and become more integrated into educational practice.

These limitations suggest productive directions for future research. First, cross-cultural validation studies are needed to examine the scale's performance in different educational systems and cultural contexts. Second, longitudinal studies could track how teachers' attitudes evolve as they gain more experience with AI

ISSN : 2410-1818

tools, potentially revealing patterns of acceptance development. Third, research exploring the relationship between attitude levels measured by this scale and actual classroom implementation of AI would strengthen its predictive validity. Finally, qualitative investigations could provide deeper insight into the psychological and contextual factors that shape the critical transition points identified in the hierarchical scale, particularly the shift toward accepting AI's role in transformative practices and potential teacher replacement.

ISSN : 2410-1818

References

Afif (2023). Applications of artificial intelligence in education. Planning and Development Department, Afif Education Administration.

Al-Dulaimi, N. K. J., & Al-Amiri, A. H. M. (2024). The effect of AI usage on achievement and attitude toward history topics among middle school students. *Middle East Journal of Educational Sciences*, 1828–2789.

Al-Muslim, N. I. I. (2023). Science teachers' attitudes toward using AI applications in elementary education in Jazan Education Administration. Jazan University, Saudi Arabia.

Al-Shammari, S. (2024). Teachers' attitudes toward employing AI applications to address learning difficulties in Asir, Saudi Arabia. *Journal of Reading and Knowledge*, 263, 95–134.

Alsoudi, S., & Abu Shindi, Y. (2023). Rasch Rating Scale Modeling of the Arabic Version of the Critical Thinking Disposition Scale. *Journal of Educational and Psychological Studies Sultan Qaboos University*, 17(4), 359–369. https://doi.org/10.24200/jeps.2023.56789

Alsoudi, S., & AlHarthy, S. (2024). Applying the Guttman method and Rasch model to construct school students' attitudes toward the significance of the mathematics scale. *Journal of Southwest Jiaotong University*, 59(3), 488–499. https://doi.org/10.35741/issn.0258-2724.59.3.488

Bond, M., Zawacki-Richter, O., & Nichols, M. (2020). Revisiting five decades of educational technology research: A content and authorship analysis of the British Journal of Educational Technology. *British Journal of Educational Technology, 50*(1), 12-63. https://doi.org/10.1111/bjet.12890

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230-258. https://doi.org/10.1177/0049124192021002005

Çayak, A. (2024). Teachers' attitudes toward the integration of AI in classroom environments: A systematic review. *Journal of Educational Technology Studies*, 12(1), 15–38.

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *Educational Research Review*, 36, 100447. https://doi.org/10.1016/j.edurev.2022.100447

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual

ISSN : 2410-1818

correlations. *Applied Psychological Measurement, 41*(3), 178–194. https://doi.org/10.1177/0146621616677520

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Fisher, W. P. (1992). Reliability statistics. *Rasch Measurement Transactions, 6*(3), 238.

Ghamdi, M. F. M. (2024). Artificial intelligence in education. *King Fahd National Library.*

Granić, A., & Marangunić, N. (2023). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology, 50*(5), 2570-2593. https://doi.org/10.1111/bjet.13330

Guttman, L. (1950). The basis for scalogram analysis. In S. A. Stouffer (Ed.), *Measurement and prediction* (pp. 60-90). Princeton University Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Ibrahim, E. O. (2024). Employing AI to develop academic programs: A descriptive-analytical study among faculty at Hamad Bin Khalifa University. *Proceedings of the 27th Annual Conference of Specialized Libraries Association, Gulf Branch*, 1119–1136.

Jisc. (2021). *The future of assessment: Five principles, five targets for 2025*. Jisc. https://www.jisc.ac.uk/reports/the-future-of-assessment

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.

Linacre, J. M. (2024). *A user's guide to WINSTEPS® MINISTEP Rasch-model computer programs*. https://www.winsteps.com/a/winsteps-manual.pdf

Loevinger, J. (1948). The technique of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychological Bulletin, 45*(6), 507-530.

Ludlow, L. H., Baez-Cruz, C., Chang, W., & Reynolds, M. (2025). Hierarchical scaling and Rasch measurement in educational research. *Educational Measurement: Issues and Practice*, 44(1), 12–27. https://doi.org/10.1111/emip.12456

Müller, M. (2023). A comprehensive guide to Guttman scaling: Theory and applications. *Educational Measurement Review, 45*(2), 123-145.

ISSN : 2410-1818

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Raîche, G. (2005). Critical eigenvalue sizes in standardized residual principal components analysis. *Rasch Measurement Transactions, 19*(1), 1012-1013. https://www.rasch.org/rmt/rmt191a.htm

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*(3), 207-230. https://doi.org/10.3102/10769986004003207

Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education, 128*, 13-35. https://doi.org/10.1016/j.compedu.2018.09.009

Smith, R. M., Miao, C. Y., & Sa, B. (1994). Assessing unidimensionality for Rasch measurement. In M. Wilson (Ed.), *Objective measurement: Theory into practice* (Vol. 2, pp. 316-327). Ablex Publishing.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55.

UNESCO. (2023). Artificial intelligence in education: Challenges and opportunities. Paris: UNESCO Publishing. https://doi.org/10.1787/9789264505132-en

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Wang, W.-C., Chen, H.-F., & Jin, K.-Y. (2018). Item response models for local dependence and testlet effects. In *Handbook of item response theory modeling* (pp. 183–206). Routledge.

Wang, X., Li, B., & Wang, J. (2023). Teacher technostress and coping strategies during COVID-19: A systematic review. *Educational Technology Research and Development, 71*(1), 1–25. https://doi.org/10.1007/s11423-023-10203-6

Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, 126, 103505. https://doi.org/10.1016/j.system.2024.103505

Wright, B. D., & Linacre, J. M. (1989). Combining and splitting categories. *Rasch Measurement Transactions, 3*(2), 59–60.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education

ISSN : 2410-1818

– where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, K., Wu, H., & Asare, S. K. (2023). Understanding AI acceptance in education: A unified model based on the technology acceptance framework and attitude theory. *Computers in Human Behavior, 148*, 107903. https://doi.org/10.1016/j.chb.2023.107903

Zheng, L., Niu, J., Long, M., & Fan, Y. (2022). The ceiling effect in educational technology research: Diagnosis and treatment. *Journal of Educational Computing Research, 60*(4), 1001-1025. https://doi.org/10.1177/07356331211045339

ISSN : 2410-1818

مجلـــة الأندلـــس للعلـــوم الإنسانيـــة والاجتماعيـــة
مجلـــة دوليـــة شهـريـــة علميـــة محكمـــة
الترقيـــم الدولي الإلكتروني : ISSN 2410- 521X
الترقيـــم الدولـــي الورقـــي :ISSN 2410- 1818
البريد الإلكتروني : journal@andalusuniv.net

المجلـــة مفهرســـة فـــي المواقـــع الآتيـــة :

العام	2022	2023	2024	2025	2026
معامل أرسيف	0.1954	0.3759	0.3068	0.5978	0.4135
معامل التأثير العربي	1.73	1.25	1.55	1.59	1.58