BIO-INOVED
Jurnal Biologi-Inovasi Pendidikan
Volume 8, Issue 1, February 2026 page. 112-122

P-ISSN : 2684-9062 | E-ISSN : 2714-9803
https://ppjp.ulm.ac.id/journal/index.php/bino
DOI: 10.20527/bino.v8i1.24953

Rasch model–based development of an environmental literacy instrument

Moch. Dwi Irsyad Saputra [(1)] *, Endang Susilaningsih [(2)], Ani Rusilowati [(2)], Yuniar Fahmi Lathif [(3)]

[(1)] Doctoral Program of Science Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Semarang, Central Java, Indonesia
[(2)] Department of Science Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Semarang, Central Java, Indonesia
[(3)] Master's Program in Primary Islamic Teacher Education, Postgraduate Program, Universitas Islam Negeri Salatiga, Salatiga, Central Java, Indonesia

** Corresponding Author Email: irsyadamanda@students.unnes.ac.id*

Article Information

Keyword:
Rasch model; Environmental literacy; Environmental behavior; Psychometric properties

Kata Kunci:
Model Rasch; Literasi lingkungan; Perilaku lingkungan; Instrumen psikometri

History:
Received : 22/12/2025
Revised : 16/02/2026
Accepted : 20/02/2026
Published : 22/02/2026

Abstract

This study aims to develop and psychometrically validate an innovative environmental literacy instrument by integrating contextualized behavioral indicators using the Rasch model, offering a novel approach compared to existing instrument. The instrument consisted of 20 items on a five-point Likert scale and was administered to 124 junior secondary school students. Data were analyzed using Winsteps 3.73, focusing on reliability, separation indices, item fit, unidimensionality, and item difficulty. Person analysis showed a reliability of 0.94 with a separation index ranging from 3.87 to 4.03, indicating the instrument can distinguish respondents into approximately four behavioral levels. Item analysis produced an item reliability of 0.98 with separation indices between 7.13 and 7.20, reflecting a stable hierarchy of item difficulty. All items fit the Rasch Model, with Infit Mean Square values between 0.81 and 1.12 and Outfit Mean Square values from 0.81 to 1.16. Positive PT-Measure Correlations ranging from 0.49 to 0.78 confirmed alignment with the measured construct. Unidimensionality analysis revealed that 60.8% of variance was explained by the model, while the first contrast accounted for 2.0 eigenvalues (4.0%), supporting a single dominant construct. Item difficulty estimates ranged from –1.58 to 1.53 logits (SD = 0.87). Overall, the findings confirm that the instrument is valid, reliable and suitable for assessing students' environmental literacy behavior in educational research and evaluation contexts, with potential applications in curriculum evaluation and environmental education program assessment.

Abstrak. *Penelitian ini bertujuan untuk mengembangkan dan memvalidasi secara psikometrik instrumen literasi lingkungan yang inovatif dengan mengintegrasikan indikator perilaku terkkontekstualisasi menggunakan model Rasch, sehingga menawarkan pendekatan baru dibandingkan instrumen yang telah ada. Instrumen terdiri atas 20 butir dengan skala Likert lima tingkat dan diujikan kepada 124 siswa sekolah menengah pertama. Analisis data dilakukan menggunakan perangkat lunak Winsteps 3.73 dengan fokus pada reliabilitas, indeks separasi, unidimensionalitas konstruk dan tingkat kesulitan butir instrumen. Hasil analisis menunjukkan bahwa analisis kemampuan responden menunjukkan reliabilitas person sebesar 0,94 dengan indeks separasi 3,87–4,03, yang menandakan kemampuan instrumen dalam membedakan responden ke dalam sekitar empat tingkat perilaku literasi lingkungan. Pada sisi kualitas butir, diperoleh reliabilitas item sebesar 0,98 dengan indeks separasi 7,13–7,20, yang menunjukkan stabilitas hierarki tingkat kesulitan butir. Seluruh butir instrumen dinyatakan fit terhadap Model Rasch dengan nilai Infit Mean Square berkisar antara 0,81–1,12 dan Outfit Mean Square antara 0,81–1,16, serta korelasi butir–konstruk (PT-Measure Correlation) yang positif (0,49–0,78). Uji unidimensionalitas melalui analisis residual terstandar menunjukkan varians yang dijelaskan oleh model sebesar 60,8%, dengan unexplained variance pada kontras pertama sebesar 2,0 eigenvalue (4,0%), yang mengindikasikan bahwa instrumen mengukur satu konstruk utama. Nilai kesulitan item tersebar dari –1,58 hingga 1,53 logit (SD = 0,87), dengan distribusi 4 butir sangat mudah, 7 butir mudah, 5 butir sulit, dan 4 butir sangat sulit. Secara keseluruhan, temuan ini menegaskan bahwa penggunaan Model Rasch memberikan informasi diagnostik yang komprehensif dan akurat terhadap kualitas instrumen, sehingga instrumen dinyatakan layak dan andal untuk mengukur perilaku literasi lingkungan peserta didik dalam konteks penelitian dan evaluasi pembelajaran, serta dimanfaatkan untuk evaluasi kurikulum dan penilaian program pendidikan lingkungan.*

© 2026 BIO-INOVED : Jurnal Biologi-Inovasi Pendidikan

A. Introduction

Global attention to environmental issues has increased; however, students' pro-environmental behavior remains uneven across educational contexts. Most existing environmental literacy instruments primarily emphasize cognitive and attitudinal dimensions, while behavioral aspects are insufficiently operationalized and measured. This limitation underscores the need for a behavior-focused environmental literacy instrument capable of providing robust evidence for educational evaluation. ,. Numerous studies indicate that many students and university learners still exhibit low levels of scientific and environmental literacy, both in terms of conceptual understanding and the application of knowledge in real-world contexts (Adhari et al., 2025; Teksoz et al., 2014). Environmental literacy frameworks generally encompass ecological knowledge, attitudes, cognitive skills, and pro-environmental behavior, which are viewed as a complex and multidimensional construct. Consequently, this condition demands measurement instruments that are concise yet remain valid and reliable (Szczytko et al., 2018). High-quality instruments are required to support systematic and comparable evaluations of environmental education programs across formal and nonformal contexts (Markos et al., 2017).

However, the measurement of constructs in education,such as literacy, digital literacy, health literacy, and pro-environmental attitudes still predominantly relies on Classical Test Theory (CTT) and raw score approaches. These limitations of Classical Test Theory are especially critical for environmental behavior constructs, which are inherently context-dependent and multidimensional. Pro-environmental behaviors vary across situations and social settings, making score interpretation highly sensitive to sample characteristics and item composition. Consequently, CTT-based instruments may fail to capture stable behavioral patterns, underscoring the need for Rasch-based measurement approaches that provide item- and person-invariant estimates. These methods are highly sample-dependent, limiting cross-group comparability, reducing generalizability, and constraining the ability to simultaneously represent respondent ability and item difficulty (Avinç & Doğan, 2024; Neumann et al., 2011; Chan et al., 2020). Previous instrument development studies have shown that without rigorous psychometric analysis, instruments may contain misfitting or biased items and may fail to discriminate respondents across different ability levels (Erfan et al., 2020; Petra & Aziz, 2020). Such limitations can lead to biased ability estimates when instruments are applied to heterogeneous populations.

In educational practice, evaluation instruments used in schools often emphasize cognitive mastery of concepts while insufficiently integrating authentic environmental issues or attitudinal and behavioral dimensions (Adhari et al., 2025; Sukri et al., 2022). Environmental literacy is not solely knowledge-oriented but also involves attitudes, critical thinking skills, and observable behaviors. Within modern measurement frameworks, these dimensions are conceptualized as latent traits that must be estimated with precision. Recent trends in environmental literacy and attitude instrument development highlight the increasing use of Item Response Theory, particularly the Rasch Model, to ensure that each item meaningfully contributes to the intended construct, satisfies unidimensionality and local independence assumptions, and remains free from bias (Ibrahim, 2025; Avcu, 2025).

The Rasch Model is regarded as an appropriate approach because it calibrates respondent ability and item difficulty on the same interval scale (logit), enables person–item mapping, and provides detailed information on person and item reliability, response category functioning, and differential item functioning (DIF) detection (Avcu, 2025; Erfan et al., 2020). Empirical studies on scale development in education and health—such as digital literacy scales (Avinç & Doğan, 2024), adolescent health literacy (Fleary et al., 2021), environmental health literacy among health professionals (Vrotsou et al., 2023), pro-environmental attitudes (Barón & Baez, 2022), and children's environmental health knowledge and skills (Şimşek & Dağhan, 2025)—demonstrate that Rasch-based instruments yield more accurate, robust, and sustainable measurement outcomes across diverse educational contexts. This advantage arises from the Rasch Model's capacity to clarify construct structure, enhance item quality, and strengthen measurement reliability for data-driven educational decision-making (Supriatna et al., 2025; Qudratuddarsi et al., 2022).

Although numerous environmental literacy instruments have been developed, most primarily emphasize cognitive and attitudinal dimensions, with limited attention to behavioral aspects, particularly those examined through Rasch analysis. This gap highlights the need for a behavior-focused environmental literacy instrument grounded in Rasch-based measurement to provide more robust and invariant estimates of students' pro-environmental behavior. Therefore, the development of environmental literacy instruments grounded in the Rasch Model is an urgent necessity to obtain a comprehensive, calibrated, and comparable representation of students' environmental literacy competencies across contexts. Such instruments will strengthen environmental education program evaluation, support improvements in instructional practice, and provide a strong empirical foundation for sustainable environmental education policy formulation.

B. Material and method

This study employed a research and development (R&D) approach aimed at producing a Rasch Model–based environmental literacy instrument that is valid and reliable. The development model applied was ADDIE (Analysis, Design, Development, Implementation, and Evaluation), which has been widely recommended for the development of educational materials and assessment instruments (Nesri & Kristanto, 2020; Latip, 2022). Although ADDIE is commonly applied in learning material development, its systematic and iterative stages—analysis, design, development, implementation, and evaluation—also provide a structured framework for instrument development. In this study, ADDIE facilitated needs analysis, item construction, expert validation, field testing, and iterative refinement, ensuring the development of a psychometrically sound environmental literacy instrument (see Figure 1). The ADDIE model was selected due to its systematic yet flexible structure in educational evaluation development, while the Rasch Model was employed to empirically ensure the psychometric quality of the instrument (Branch, 2009; Bond & Fox, 2015).

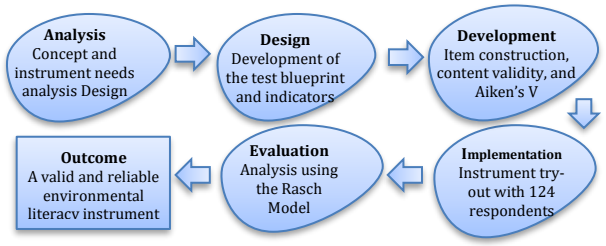

Figure 1
Stages of environmental literacy instrument development using the ADDIE model

The development stage involved constructing instrument items based on a predefined blueprint and conducting content validity through expert judgment by specialists in science education and educational assessment. Content validity was analyzed using Aiken's V index, with items considered acceptable when V > 0.62 (Retnawati, 2016; Azwar, 2018). The items were revised according to expert feedback until a trial-ready instrument was obtained.

The implementation stage consisted of pilot testing the instrument with research participants. The population included all students at the targeted educational level at MTs Tahfidz Yanbu'ul Qur'an Kudus. The sample was determined using convenience sampling, resulting in 124 students. This sample size meets the minimum recommendation for Rasch analysis, namely ≥100 respondents for polytomous data to ensure stable parameter estimation (Putri & Aini, 2023). Data were analyzed using the Rasch Rating Scale Model in Winsteps 3.73. The use of this software aligns with current practices in educational and literacy instrument development that emphasize the integration of content validity and item response modeling (Firdaus et al., 2025; Kania et al., 2024).

The evaluation stage involved Rasch Model analysis to assess the psychometric quality of the instrument (Boone, 2016; Rasyid et al., 2025). The analysis results served as the basis for revising or eliminating misfitting items, resulting in a final environmental literacy instrument that is valid, reliable, and demonstrates high measurement precision (Ibrahim, 2025; Permatasari, 2025).

C. Results and discussion

At the analysis stage, a theoretical review of environmental literacy was conducted based on the framework proposed by the North American Association for Environmental Education (NAAEE), which comprises four core dimensions: environmental knowledge, environmental attitudes, cognitive skills, and environmental behavior (NAAEE, 2011; McBeth & Volk, 2009). Although the NAAEE framework encompasses multiple dimensions of environmental literacy, this study intentionally focused on the behavioral dimension. Environmental behavior represents the most observable and actionable outcome of environmental literacy and provides direct evidence of students' engagement in pro-environmental practices. Moreover, behavioral indicators are particularly suitable for Rasch-based measurement because they can be operationalized as context-specific actions, allowing more precise estimation of students' environmental literacy levels. Other NAAEE dimensions were not included to maintain construct focus and to ensure robust psychometric evaluation within the scope of this study. In addition, an instrument needs analysis was carried out through a review of evaluation practices in schools, which revealed that the environmental literacy instruments currently in use remain conventional and have not been subjected to rigorous psychometric analysis. This finding aligns with the ADDIE model principles emphasizing systematic analysis of instructional goals, content, and assessment needs.

The design stage involved developing an instrument blueprint based on environmental literacy indicators, constructing the questionnaire items, and determining the measurement scale. The instrument was designed using a five-point Likert scale to facilitate polytomous analysis within the Rasch Model framework. At this stage, scoring procedures and the instrument validation plan were also established (Bashooir & Supahar, 2018; Nurfadillah & Elvia, 2023).

The initial development phase of the instrument included content validation conducted by four experts with expertise in environmental education and educational assessment. This validation aimed to evaluate the alignment of questionnaire items with environmental behavior indicators, the clarity of item wording, and the overall representativeness of the

measured construct (Azwar, 2018). The environmental behavior indicators assessed included waste management habits, participation in environmental conservation activities, environmentally friendly lifestyles, and concern for natural resource conservation (Hungerford & Volk, 1990; NAAEE, 2010).

Table 1 Results of content validity analysis using Aiken's V index

Aspect	Description
Number of instrument items	20 items
Instrument type	Five-point Likert-scale questionnaire
Measured construct	Environmental Behavior
Number of expert validators	4 experts
Range of Aiken's V values	0.63–0.94
Mean Aiken's V value	0.82
Minimum acceptance criterion	V ≥ 0.62
Number of valid items	20 items (100%)

Table 2 Person and item reliability and separation index analysis

INPUT: 124 PERSON 20 ITEM REPORTED: 124 PERSON 20 ITEM 5 CATS WINSTEPS 3.7

SUMMARY OF 124 MEASURED PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	67.0	20.0	.56	.30	.96	-.1	.98	.0
S.D.	14.9	.0	1.25	.03	.24	.8	.25	.8
MAX.	91.0	20.0	2.92	.40	1.56	1.7	1.77	2.1
MIN.	31.0	20.0	-2.55	.27	.42	-2.4	.41	-2.5

REAL RMSE	.31	TRUE SD	1.21	SEPARATION 3.87	PERSON RELIABILITY .94
MODEL RMSE	.30	TRUE SD	1.22	SEPARATION 4.03	PERSON RELIABILITY .94
S.E. OF PERSON MEAN = .11					

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .95

SUMMARY OF 20 MEASURED ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	415.6	124.0	.00	.12	.95	-.4	.98	-.2
S.D.	63.0	.0	.87	.01	.08	.7	.10	.8
MAX.	521.0	124.0	1.53	.14	1.12	.9	1.16	1.1
MIN.	301.0	124.0	-1.58	.11	.81	-1.6	.81	-1.6

REAL RMSE	.12	TRUE SD	.86	SEPARATION 7.13	ITEM RELIABILITY .98
MODEL RMSE	.12	TRUE SD	.86	SEPARATION 7.20	ITEM RELIABILITY .98
S.E. OF ITEM MEAN = .20					

Based on the Aiken's V analysis presented in Table 1, all 20 questionnaire items obtained V values greater than 0.62 and therefore met the criteria for adequate content validity (Aiken, 1985; Azwar, 2018). Although the Aiken's V coefficients indicated strong content validity, very high values may also suggest potential item redundancy. To address this concern, items were carefully reviewed for conceptual overlap, and redundant indicators were refined or removed to preserve construct coverage while minimizing duplication. The Aiken's V coefficients ranged from 0.63 to 0.94, indicating good to excellent agreement between the instrument items and the environmental behavior indicators being measured. Items with higher

Aiken's V values (≥ 0.82) reflect strong consensus among experts regarding the relevance of the items in accurately assessing students' environmental behavior (Azwar, 2018). These results demonstrate that the developed instrument satisfies the principle of content representativeness, ensuring comprehensive coverage of environmental behavior indicators across the questionnaire (Haynes et al., 1995). With content validity established, the instrument was deemed suitable to proceed to the implementation stage of the ADDIE model, namely empirical testing using Rasch Model analysis.

Rasch Model analysis was employed to evaluate the psychometric quality of the instrument (Boone, 2016; Rasyid et al., 2025). The analysis focused on the following aspects:

1) Person and item reliability and separation indices,
2) Item fit based on Infit and Outfit Mean Square (MNSQ) statistics,
3) Unidimensionality testing to ensure that the instrument measures a single environmental literacy construct, and
4) Item difficulty distribution to examine the spread of items from easy to difficult levels.

The results of the Rasch Model analysis presented in Table 2 indicate that the environmental literacy behavior instrument demonstrates excellent psychometric quality. The mean person measure was 0.56 logits with a standard deviation of 1.25 logits, and the ability estimates ranged from –2.55 to 2.92 logits. This distribution suggests that respondent abilities were reasonably spread without extreme clustering, as expected in Rasch-based measurement of attitudinal and behavioral traits (Tesio et al., 2023). This pattern reflects a realistic empirical condition among students, showing variability in environmental literacy behavior without excessive dominance of respondents at extremely low or high ability levels. The person reliability value of 0.94, accompanied by a separation index ranging from 3.87 to 4.03, indicates that the instrument is capable of consistently distinguishing students into several distinct ability groups, in line with recommended separation criteria in educational Rasch literature (Khalid et al., 2023; Mujiansyah et al., 2025).

Item-level analysis further confirms the robustness of the instrument. An item reliability of 0.98 indicates a highly stable hierarchy of item difficulty. The item separation index, ranging from 7.13 to 7.20, shows that the items are well distributed across a broad range of difficulty levels, from relatively easy to more demanding items. This high separation suggests that the instrument effectively represents variations in environmental literacy behavior comprehensively, ranging from basic behaviors to those requiring higher levels of commitment (Adhari et al., 2025; Supriyadi et al., 2025). The mean item difficulty was centered at 0.00 logits with a standard deviation of 0.87 logits, and difficulty estimates ranged from –1.58 to 1.53 logits. This range indicates moderate item difficulty without extreme values, confirming that all items remain relevant for measuring respondent ability. In addition, the Infit (0.96) and Outfit Mean Square (MNSQ) values (0.98) for both persons and items fall within the acceptable range (0.5–1.5), indicating that the instrument fits the Rasch Model well.

The results of the response category structure analysis presented in Table 3 indicate that all five response categories of the Likert scale functioned effectively. The average measures increased monotonically across categories, indicating that higher response categories corresponded to higher respondent ability levels. The Infit and Outfit Mean Square (MNSQ) values for all categories fell within the acceptable range (0.5–1.5), and the Andrich thresholds were ordered sequentially without any evidence of disordered thresholds. These findings suggest that respondents were able to clearly differentiate among the response categories, and therefore the response scale is considered valid and does not require category collapsing (Touzani et al., 2024; Van Zile-Tamsen, 2017).

Table 3 Analysis of Likert scale response category structure

INPUT: 124 PERSON 20 ITEM REPORTED: 124 PERSON 20 ITEM 5 CATS WINSTEPS 3.73

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE	
1	1	188	8	-1.51	-1.66	1.08	1.05	NONE	(-3.31)	1
2	2	435	18	-.89	-.77	.90	.92	-2.05	-1.49	2
3	3	621	25	.07	.17	.86	.83	-.66	-.08	3
4	4	790	32	1.39	1.19	.86	1.03	.43	1.47	4
5	5	446	18	2.06	2.23	1.08	1.07	2.28	(3.48)	5

Table 4 Item fit analysis based on the Rasch model

INPUT: 124 PERSON 20 ITEM REPORTED: 124 PERSON 20 ITEM 5 CATS WINSTEPS 3.73

PERSON: REAL SEP.: 3.87 REL.: .94 ... ITEM: REAL SEP.: 7.13 REL.: .98

ITEM STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	IQIAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
1	521	124	-1.58	.14	.89	-.7	1.16	1.1	A .49	.63	55.6	58.2	I01
20	496	124	-1.13	.13	1.12	.9	1.13	1.0	B .59	.66	42.7	53.9	I20
19	301	124	1.53	.12	1.11	.9	1.11	.9	C .69	.70	46.8	47.6	I19
11	473	124	-.77	.12	1.01	.1	1.08	.7	D .63	.68	43.5	50.9	I11
9	361	124	.75	.11	1.08	.7	1.07	.6	E .73	.72	39.5	45.9	I09
3	495	124	-1.12	.13	1.00	.0	1.05	.4	F .68	.66	58.1	53.8	I03
13	425	124	-.09	.12	1.04	.4	1.01	.1	G .71	.71	47.6	47.0	I13
12	466	124	-.66	.12	.99	.0	1.02	.2	H .71	.68	47.6	50.0	I12
16	350	124	.89	.11	1.00	.1	1.01	.2	I .68	.72	40.3	45.8	I16
14	388	124	.40	.11	.94	-.4	1.00	.1	J .74	.71	46.0	46.5	I14
2	490	124	-1.04	.13	.88	-.9	.98	-.1	j .63	.67	47.6	52.9	I02
17	440	124	-.29	.12	.96	-.3	.94	-.5	i .71	.70	43.5	47.7	I17
7	397	124	.28	.11	.95	-.4	.93	-.5	h .75	.71	45.2	46.7	I07
5	436	124	-.24	.12	.92	-.6	.92	-.6	g .74	.70	48.4	47.6	I05
6	432	124	-.18	.12	.91	-.7	.91	-.7	f .74	.70	44.4	47.6	I06
15	377	124	.54	.11	.86	-1.1	.89	-.9	e .76	.72	46.8	46.1	I15
8	371	124	.62	.11	.89	-.9	.86	-1.1	d .78	.72	46.0	46.0	I08
4	444	124	-.35	.12	.87	-1.0	.85	-1.3	c .70	.70	55.6	47.8	I04
10	330	124	1.15	.11	.86	-1.2	.86	-1.2	b .74	.71	46.8	46.2	I10
18	318	124	1.30	.11	.81	-1.6	.81	-1.6	a .71	.71	53.2	47.1	I18
MEAN	415.6	124.0	.00	.12	.95	-.4	.98	-.2			47.3	48.8	
S.D.	63.0	.0	.87	.01	.08	.7	.10	.8			4.8	3.4	

Item fit analysis based on the Rasch Model in Table 4 was conducted by examining the Infit Mean Square (Infit MNSQ) and Outfit Mean Square (Outfit MNSQ) values for each item. This analysis aimed to ensure that each item consistently measured the same underlying construct and did not statistically deviate from the response patterns expected by the Rasch Model. Commonly accepted criteria in Rasch-based studies indicate that Infit and Outfit MNSQ values within the range of 0.5–1.5 suggest that items fit the model and function in accordance with the intended construct. This criterion is consistent with international guidelines that recommend cut-off ranges of 0.5–1.5 or 0.7–1.3 as indicators of acceptable item fit (Cantó-Cerdán et al., 2021; Park et al., 2021).

The results indicate that all instrument items exhibited Infit and Outfit MNSQ values within the acceptable range, with mean Infit MNSQ and Outfit MNSQ values of 0.95 and 0.98, respectively. Good item fit indicates that each item functions as intended, thereby enhancing the accuracy and interpretability of environmental behavior measurement in school contexts. The highest observed Outfit MNSQ value was 1.16, which remains well below the critical threshold of 1.5. This finding indicates the absence of items that measure unintended constructs or generate systematically unexpected response patterns (Smith et al., 1998; Tran et al., 2018). Consequently, all items functioned productively for measurement purposes and did not require elimination based on item fit criteria.

Furthermore, the Point–Measure Correlation (PT-Measure Corr) values for all items were positive, ranging from 0.49 to 0.78, indicating that each item contributed positively to the measured construct and demonstrated adequate discriminative power (Müller, 2020). These results confirm that the instrument items function consistently, do not exhibit random response patterns, and do not require removal based on Rasch model fit criteria. Therefore, all items are considered internally valid and support the accurate measurement of environmental literacy behavior.

The unidimensionality test presented in Table 5 was conducted using Principal Component Analysis (PCA) of standardized residuals to ensure that the instrument measures only a single dominant construct, namely environmental literacy behavior. The PCA results indicate that the instrument satisfies the unidimensionality assumption. The variance explained by the Rasch model reached 60.8%, suggesting that the majority of data variability is accounted for by one primary latent dimension. Meanwhile, the eigenvalue of the first contrast was 2.0 (4.0%), which falls within acceptable tolerance limits and does not indicate the presence of a dominant secondary latent dimension. Subsequent contrasts exhibited progressively smaller

eigenvalues and a decreasing pattern, indicating that the residual variance is random in nature and does not form additional meaningful constructs. Therefore, it can be concluded that all instrument items collectively measure a single underlying construct, namely environmental literacy behavior (Khalid et al., 2023; Ngandoh et al., 2025).

Table 5 Unidimensionality test analysis

INPUT: 124 PERSON 20 ITEM REPORTED: 124 PERSON 20 ITEM 5 CATS WINSTEPS 3.73

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	51.0	100.0%	100.0%
Raw variance explained by measures	=	31.0	60.8%	58.9%
Raw variance explained by persons	=	15.6	30.5%	29.6%
Raw Variance explained by items	=	15.4	30.2%	29.3%
Raw unexplained variance (total)	=	20.0	39.2% 100.0%	41.1%
Unexplned variance in 1st contrast	=	2.0	4.0% 10.1%	
Unexplned variance in 2nd contrast	=	1.7	3.3% 8.5%	
Unexplned variance in 3rd contrast	=	1.6	3.1% 7.9%	
Unexplned variance in 4th contrast	=	1.5	2.9% 7.4%	
Unexplned variance in 5th contrast	=	1.4	2.7% 6.8%	

Table 6 Analysis of instrument item difficulty levels

INPUT: 124 PERSON 20 ITEM REPORTED: 124 PERSON 20 ITEM 5 CATS WINSTEPS 3.73

PERSON: REAL SEP.: 3.87 REL.: .94 ... ITEM: REAL SEP.: 7.13 REL.: .98

ITEM STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
19	301	124	1.53	.12	1.11	.9	1.11	.9	.69	.70	46.8	47.6	I19
18	318	124	1.30	.11	.81	-1.6	.81	-1.6	.71	.71	53.2	47.1	I18
10	330	124	1.15	.11	.86	-1.2	.86	-1.2	.74	.71	46.8	46.2	I10
16	350	124	.89	.11	1.00	.1	1.01	.2	.68	.72	40.3	45.8	I16
9	361	124	.75	.11	1.08	.7	1.07	.6	.73	.72	39.5	45.9	I09
8	371	124	.62	.11	.89	-.9	.86	-1.1	.78	.72	46.0	46.0	I08
15	377	124	.54	.11	.86	-1.1	.89	-.9	.76	.72	46.8	46.1	I15
14	388	124	.40	.11	.94	-.4	1.00	.1	.74	.71	46.0	46.5	I14
7	397	124	.28	.11	.95	-.4	.93	-.5	.75	.71	45.2	46.7	I07
13	425	124	-.09	.12	1.04	.4	1.01	.1	.71	.71	47.6	47.0	I13
6	432	124	-.18	.12	.91	-.7	.91	-.7	.74	.70	44.4	47.6	I06
5	436	124	-.24	.12	.92	-.6	.92	-.6	.74	.70	48.4	47.6	I05
17	440	124	-.29	.12	.96	-.3	.94	-.5	.71	.70	43.5	47.7	I17
4	444	124	-.35	.12	.87	-1.0	.85	-1.3	.70	.70	55.6	47.8	I04
12	466	124	-.66	.12	.99	.0	1.02	.2	.71	.68	47.6	50.0	I12
11	473	124	-.77	.12	1.01	.1	1.08	.7	.63	.68	43.5	50.9	I11
2	490	124	-1.04	.13	.88	-.9	.98	-.1	.63	.67	47.6	52.9	I02
3	495	124	-1.12	.13	1.00	.0	1.05	.4	.68	.66	58.1	53.8	I03
20	496	124	-1.13	.13	1.12	.9	1.13	1.0	.59	.66	42.7	53.9	I20
1	521	124	-1.58	.14	.89	-.7	1.16	1.1	.49	.63	55.6	58.2	I01
MEAN	415.6	124.0	.00	.12	.95	-.4	.98	-.2			47.3	48.8	
S.D.	63.0	.0	.87	.01	.08	.7	.10	.8			4.8	3.4	

Table 7 Categories of instrument item difficulty

No.	Difficulty category	Logit range	Number of items
1	Very difficult	> +0.87	4
2	Difficult	0.00 to +0.87	5
3	Easy	–0.87 to < 0.00	7
4	Very easy	< –0.87	4
Total			**20**

Table 6 presents the analysis of instrument item difficulty levels. Item difficulty was examined based on item measure (logit) values derived from the Rasch Model, with the mean item difficulty centered at 0.00 logits and a standard deviation of 0.87 logits. Based on the classification of item difficulty using the mean ± standard deviation criterion commonly applied in Rasch analysis, four levels of item difficulty were identified: very difficult, difficult, easy, and very easy (Natsir et al., 2022). This classification approach has been widely used in educational test and instrument evaluation to provide an informative and balanced representation of item difficulty distribution (Octaviana et al., 2022).

The results indicate that four items were classified as very difficult, five as difficult, seven as easy, and four as very easy, as summarized in Table 7. This distribution reflects a relatively balanced spread of item difficulty, with most items located in the easy-to-difficult range, which is considered optimal for measuring environmental literacy behavior. Very easy items represent basic behaviors commonly practiced by students and function to identify respondents at lower to moderate ability levels. In contrast, very difficult items reflect environmental literacy behaviors that require higher levels of awareness, consistency, and commitment, similar to findings reported in instruments assessing higher-order thinking skills or complex pro-environmental attitudes (Prasetya & Pratama, 2023).

The relatively balanced distribution, with the majority of items concentrated in the easy-to-difficult categories, indicates that the instrument is well targeted to the ability range of most respondents while still including anchor items at both lower and upper extremes. This pattern aligns with recommendations for Rasch-based test design, which emphasize the importance of matching item difficulty distribution with respondent ability distribution to maximize measurement information across the ability continuum (Giguère et al., 2022; Yang et al., 2025). Moreover, the presence of all four difficulty categories strengthens the instrument's capacity to sensitively capture variations in environmental literacy behavior, ranging from basic actions to more advanced and sustained behaviors. This approach is consistent with item difficulty mapping practices in various educational and literacy assessment contexts, including science, mathematics, and reading literacy (Wiradika & Retnawati, 2022). Therefore, the composition and distribution of item difficulty not only satisfy Rasch technical criteria but also support the substantive purpose of the instrument in measuring environmental literacy behavior progressively and hierarchically.

Although the Rasch Model is not conceptually intended to replace Exploratory Factor Analysis (EFA) or Confirmatory Factor Analysis (CFA), it provides a distinct and independent framework for construct validation. Consistent with previous empirical studies, Rasch analysis has been shown to provide sufficient evidence for construct validity through item fit, unidimensionality, and measurement invariance, supporting its use as a standalone validation approach in educational measurement. The Rasch Model is grounded in a probabilistic measurement paradigm that evaluates the fit of empirical data to a unidimensional measurement model rather than relying on inter-item covariance patterns. It offers a comprehensive alternative for construct validation through unidimensionality testing, local independence, item fit, reliability, and separation indices (Touzani et al., 2024; Huang & Chen, 2025). Therefore, the application of EFA and CFA is not a methodological necessity when Rasch analysis has satisfied all required measurement criteria and the research objective is focused on a single latent construct (Ibrahim, 2025; Han, 2022).

D. Conclusion

Based on the Rasch Model results, the developed environmental literacy behavior questionnaire demonstrated high reliability, satisfactory item fit, and clear unidimensionality, indicating strong psychometric quality. This is evidenced by all items fitting the model, the fulfillment of the unidimensionality assumption, and very high reliability and separation indices for both persons and items. The item difficulty analysis revealed a balanced distribution of items ranging from very easy to very difficult, indicating that the instrument is capable of comprehensively measuring students' environmental literacy behavior across different ability levels. These findings confirm that the developed instrument is suitable for use as a measurement tool in both research and environmental literacy–based educational evaluation contexts; however, the results should be interpreted with caution due to the relatively limited sample size and the single-school setting of the study. The application of the Rasch Model provides significant methodological advantages over classical measurement approaches, as it allows for the independent estimation of respondent ability and item difficulty that is not dependent on sample or test characteristics. In addition, the Rasch Model offers detailed diagnostic information regarding item fit, response scale functioning, and construct structure, enabling instrument quality to be evaluated and refined in a more objective and transparent manner. Therefore, the use of the Rasch Model not only enhances measurement accuracy and reliability but also strengthens construct validity, while future studies are encouraged to conduct cross-cultural validation and differential item functioning analyses across gender and grade levels to further support the instrument's generalizability, making it highly relevant and essential for the development and evaluation of

environmental literacy behavior instruments in educational contexts.

E. References

Adhari, D., Santiani, S., & Fatmawati, S. (2025). Validity and reliability analysis of an environmentally integrated science literacy test using Rasch model. *EduFisika: Jurnal Pendidikan Fisika, 10*(1), 38-53. DOI: https://doi.org/10.59052/edufisika.v10i1.42136

Aiken, L. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement, 40*(4), 955–959. DOI: https://doi.org/10.1177/001316448004000419

Aiken, L. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131–142. DOI: https://doi.org/10.1177/0013164485451012

Avcu, A. (2025). The contribution of Rasch modeling on measuring attitudes for better classroom assessment. *International Journal of Psychology and Educational Studies, 12*(3), 169-179. DOI: https://doi.org/10.52380/ijpes.2025.12.3.1432

Avinç, E., & Doğan, F. (2024). Digital literacy scale: Validity and reliability study with the Rasch model. *Education and Information Technologies, 29*, 22895–22941. DOI: https://doi.org/10.1007/s10639-024-12662-7

Azwar, S. (2018). *Reliabilitas dan validitas* (Edisi 4). Yogyakarta: Pustaka Pelajar.

Barón, W., & Baez, A. (2022). Psychometric indicators of the pro-environmental attitudes' questionnaire: Colombian version. *Frontiers in Psychology, 13,* 886769. DOI: https://doi.org/10.3389/fpsyg.2022.886769

Bashooir, K., & Supahar, S. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(2), 219-230. DOI: https://doi.org/10.21831/pep.v22i2.19590

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

Boone, W. (2016). Rasch analysis for instrument development: Why, when, and how? *CBE Life Sciences Education, 15*(4), rm4. DOI: https://doi.org/10.1187/cbe.16-04-0148

Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.

Cantó-Cerdán, M., Cacho-Martínez, P., Lara-Lacárcel, F., & García-Muñoz, Á. (2021). Rasch analysis for development and reduction of symptom questionnaire for visual dysfunctions (SQVD). *Scientific Reports*, *11*(1), 14855. DOI: https://doi.org/10.1038/s41598-021-94166-9

Chan, S., Looi, C., & Sumintono, B. (2020). Assessing computational thinking abilities among Singapore secondary students: A Rasch model measurement analysis. *Journal of Computers in Education, 8*, 213–236. DOI: https://doi.org/10.1007/s40692-020-00177-2

Erfan, M., Maulyda, M., Ermiana, I., Hidayati, V., & Widodo, A. (2020). Validity and reliability of cognitive tests study and development of elementary curriculum using Rasch model. *Psychology, Evaluation, and Technology in Educational Research, 3*(1), 26–33. DOI: https://doi.org/10.33292/petier.v3i1.51

Firdaus, F., Wiyanto, W., Putra, N., & Isnaeni, W. (2025). Design of instruments for scientific creative thinking skills and creative thinking digital skills: Rasch models and confirmatory factor analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(5), em2632. DOI: https://doi.org/10.29333/ejmste/16310

Fleary, S., Freund, K., & Nigg, C. (2021). Development and validation of assessments of adolescent health literacy: A Rasch measurement model approach. *BMC Public Health, 22*(1), 585. DOI: https://doi.org/10.1186/s12889-022-12924-4

Giguère, G., Brouillette-Alarie, S., & Bourassa, C. (2022). A look at the difficulty and predictive validity of LS/CMI items with Rasch modeling. *Criminal Justice and Behavior, 50*, 118–138. DOI: https://doi.org/10.1177/00938548221131956

Han, H. (2022). The effectiveness of weighted least squares means and variance adjusted based fit indices in assessing local dependence of the Rasch model: Comparison with principal component analysis of residuals. *Plos one*, *17*(9), e0271992. DOI: https://doi.org/10.1371/journal.pone.0271992

Haynes, S., Richard, D., & Kubany, E. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238–247. DOI: https://doi.org/10.1037/1040-3590.7.3.238

Huang, M., & Chen, Z. (2025). Evaluating pain outcomes in Chinese ophthalmology patients using the APS-POQ-R-C: A Rasch analysis. *Frontiers in Psychology, 16*, 1558111. DOI: https://doi.org/10.3389/fpsyg.2025.1558111

Hungerford, H. R., & Volk, T. L. (1990). Changing learner behavior through environmental education. *The Journal of Environmental Education, 21*(3), 8–21. DOI: https://doi.org/10.1080/00958964.1990.10753743

Ibrahim, N. (2025). Analisis literatur penggunaan model Rasch untuk validasi instrumen asesmen pendidikan: Tren, temuan, dan implikasi. *Lencana: Jurnal Inovasi Ilmu Pendidikan*, *3*(2), 46-60. DOI: https://doi.org/10.55606/lencana.v3i2.5012

Kania, N., Kusumah, Y., Dahlan, J., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in*

Education), 10(1), 64-79. DOI: https://doi.org/10.21831/reid.v10i1.71032

Khalid, N., Yusof, I., Latif, A., & Jani, M. (2023). Psychometric properties of teacher classroom assessment literacy instrument using Rasch model analysis. *International Journal of Evaluation and Research in Education (IJERE), 12*(2), 638-646. DOI: https://doi.org/10.11591/ijere.v12i2.23607

Latip, A. (2022). Penerapan model ADDIE dalam pengembangan multimedia pembelajaran berbasis literasi sains. *DIKSAINS: Jurnal Ilmiah Pendidikan Sains, 2*(2), 102-108. DOI: https://doi.org/10.33369/diksains.2.2.102-108

Markos, A., Boubonari, T., Mogias, A., & Kevrekidis, T. (2017). Measuring ocean literacy in pre-service teachers: Psychometric properties of the Greek version of the Survey of Ocean Literacy and Experience (SOLE). *Environmental Education Research, 23*, 231–251. DOI: https://doi.org/10.1080/13504622.2015.1126807

McBeth, W., & Volk, T. L. (2009). The National Environmental Literacy Project: A baseline study of middle grade students in the United States. *The Journal of Environmental Education, 41*(1), 55–67. DOI: https://doi.org/10.1080/00958960903210031

Mujiansyah, M., Ridwan, A., & Rahmawati, Y. (2025). Psychometric analysis: Differences in person and item reliability on the adaptation of the Islamic school environment instrument in Indonesia. *International Journal of Innovative Research and Scientific Studies, 8*(3), 4556–4562. DOI: https://doi.org/10.53894/ijirss.v8i3.7548

Müller, M. (2020). Item fit statistics for Rasch analysis: Can we trust them? *Journal of Statistical Distributions and Applications, 7*(1), 5. DOI: https://doi.org/10.1186/s40488-020-00108-7

NAAEE. (2010). *Guidelines for excellence in environmental education*. North American Association for Environmental Education.

NAAEE. (2011). *Developing a framework for assessing environmental literacy*. North American Association for Environmental Education.

Natsir, I., Mayasari, D., & Munfarikhatin, A. (2022). Rasch model: Quality of test instruments to measure students' ability in algebra. In *SHS Web of Conferences* (Vol. 149, p. 01033). EDP Sciences. DOI: https://doi.org/10.1051/shsconf/202214901033

Nesri, F., & Kristanto, Y. (2020). Pengembangan modul ajar berbantuan teknologi untuk mengembangkan kecakapan abad 21 siswa. *Aksioma: Jurnal Program Studi Pendidikan Matematika, 9*(3), 480–492. DOI: https://doi.org/10.24127/ajpm.v9i3.2925

Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. *International Journal of Science Education, 33*, 1373–1405. DOI: https://doi.org/10.1080/09500693.2010.511297

Ngandoh, S., Riandi, R., Rahmat, A., & Muslim, M. (2025). Rasch model analysis of item quality for measuring conceptual understanding of the human body system. *Jurnal Inovasi Pendidikan dan Sains, 6*(1), 89-102. DOI: https://doi.org/10.51673/jips.v6i1.2459

Nurfadillah, T., & Elvia, R. (2023). Pengembangan instrumen tes kimia berbasis literasi sains untuk mengukur literasi sains siswa. *ALOTROP*, *7*(1), 44-56. DOI: https://doi.org/10.33369/alo.v7i1.28253

Octaviana, R., Anggara, M., Jamilah, R., Darmana, A., & Suyanti, R. (2022). Analisis item soal kimia SMA menggunakan Rasch model. *Orbital: Jurnal Pendidikan Kimia*, *6*(1), 26-37. DOI: https://doi.org/10.19109/ojpk.v6i1.12248

Prasetya, W., & Pratama, A. (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. DOI: https://doi.org/10.21831/pep.v27i1.58873

Park, E., Seo, H., Blair, K., & Kang, M. (2021). Rasch analysis of the Korean-Child Behavior Checklist (K-CBCL) to assess emotional and behavioral problems in children with autism spectrum disorder. *Sage Open*, *11*(3), 21582440211040790. DOI: https://doi.org/10.1177/21582440211040790

Permatasari, N. (2025). Analisis kualitas instrumen penilaian materi keanekaragaman hayati melalui tes klasik dan Rasch model. *Bio-Pedagogi: Jurnal Pembelajaran Biologi*, *14*(1), 10-26. DOI: https://doi.org/10.20961/bio-pedagogi.v14i1.88470

Petra, T., & Aziz, M. (2020). Investigating reliability and validity of student performance assessment in higher education using Rasch model. In *Journal of Physics: Conference Series* (Vol. 1529, No. 4, p. 042088). IOP Publishing. DOI: https://doi.org/10.1088/1742-6596/1529/4/042088

Putri, Y., & Aini, F. (2023). Development of a test instrument using multiple representations to assess students understanding: A Rasch model analysis. *Jurnal Pijar Mipa*, *18*(2), 183-193. DOI: https://doi.org/10.29303/jpm.v18i2.4721

Qudratuddarsi, H., Hidayat, R., Shah, R., Nasir, N., Imami, M., & Nor, R. (2022). Rasch validation of instrument measuring Gen-Z science, technology, engineering, and mathematics (STEM) application in teaching during the pandemic. *International Journal of Learning, Teaching and Educational Research*, *21*(6), 104-121. DOI: https://doi.org/10.26803/ijlter.21.6.7

Ramdani, R., Hanurawan, F., Ramli, M., Lasan, B., & Afdal, A. (2021). Development and validation of Indonesian academic resilience scale using Rasch models. *International Journal of Instruction*, *14*(1),

105-120. DOI: https://doi.org/10.29333/iji.2021.1417a

Rasyid, F., Istiyono, E., & Gunawan, C. (2025). Developing and analyzing items of a physics conceptual understanding test on wave topics for high school students using the Rasch model. *REID (Research and Evaluation in Education)*, *11*(1), 1-16. DOI: https://doi.org/10.21831/reid.v11i1.75575

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian*. Yogyakarta: Parama Publishing.

Şimşek, H., & Dağhan, Ş. (2025). Turkish validity and reliability of children's environmental health knowledge and skills questionnaires for nursing students. *BMC nursing*, *24*(1), 226. DOI: https://doi.org/10.1186/s12912-025-02846-y

Smith, R., Schumacker, R., & Bush, M. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement, 2*(1), 66–78. Retrieved from https://eric.ed.gov/?id=ED384617

Sukri, A., Arief, M., Purwanti, E., Ramdiah, S., & Lukitasari, M. (2022). Validating student's green character instrument using factor and Rasch model. *European Journal of Educational Research*, *11*(2), 859-872. DOI: https://doi.org/10.12973/eu-jer.11.2.859

Supriatna, M., Suryana, D., Anzhali, M., & Noor, A. (2025). Evaluating career self-awareness instruments for victims of violence using the Rasch model. *BMC Psychology, 13*(1), 456. DOI: https://doi.org/10.1186/s40359-025-02762-5

Supriyadi, S., Haenilah, E., & Maulina, D. (2025). Development and validation of the critical consciousness scale of environmental issues (CCS-EI): A Rasch model analysis. *Biosfer: Jurnal Pendidikan Biologi*, *18*(1), 183-198. DOI: https://doi.org/10.21009/biosferjpb.49559

Teksoz, G., Boone, J., Tuzun, O., & Oztekin, C. (2014). An evaluation of the environmental literacy of preservice teachers in Turkey through Rasch analysis. *Environmental Education Research, 20*, 202–227. DOI: https://doi.org/10.1080/13504622.2013.768604

Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 1: The "most likely" measures coming from the model. *Disability and Rehabilitation, 46*, 591–603. DOI: https://doi.org/10.1080/09638288.2023.2169771

Touzani, R., Rouquette, A., Schultz, É., Allaire, C., Carrieri, P., Mancini, J., & Hardouin, J. (2024). Psychometric validation of the French version of two scales measuring general (HLS19-Q12) and navigational (HLS19-NAV) health literacy using the Rasch model. *BMC Public Health*, *24*(1), 3079. DOI: https://doi.org/10.1186/s12889-024-20504-x

Van Zile-Tamsen, C. (2017). Using Rasch analysis to inform rating scale development. *Research in Higher Education, 58*, 922–933. DOI: https://doi.org/10.1007/s11162-017-9448-0

Vrotsou, K., Subiza-Pérez, M., Lertxundi, A., Vergara, I., Martí-Carrera, I., De Retana, L., Duo, I., & Ibarluzea, J. (2023). Environmental health knowledge of healthcare professionals: Instrument development and validation using the Rasch model. *Environmental Research*, *235*, 116582. DOI: https://doi.org/10.1016/j.envres.2023.116582

Wiradika, I., & Retnawati, H. (2022). Item characteristics analysis using the Rasch model in the development of reading literacy instruments for elementary schools students. *Progres Pendidikan*, *3*(3), 142-150. DOI: https://doi.org/10.29303/prospek.v3i3.263

Yang, H., Alanoğlu, M., Karabatak, S., & Bradley, K. (2025). Turkish adaptation of the digital literacy scale: A Rasch analysis. *International Journal of Assessment Tools in Education*, *12*(1), 93-112. DOI: https://doi.org/10.21449/ijate.1545220