https://doi.org/10.21009/JISAE

JISAE (Journal of Indonesian Student Assessment and Evaluation)

ISSN : P-ISSN: 2442-4919 | E-ISSN: 2597-8934
Website : http://journal.unj.ac.id/unj/index.php/jisae

Vol 11 No 1 (2025)

Development and Validation of a Higher-Order Thinking Skills-Based Test on Citizen Participation in Consumer Education Using Item Response Theory

Putu Diah Ari Kusumadewi[1], Universitas Negeri Yogyakarta

Sri Wening[2] Universitas Negeri Yogyakarta

Emy Budiastuti[3] Universitas Negeri Yogyakarta

Widihastuti[4] Universitas Negeri Yogyakarta

ABSTRACT

This study aims to develop and validate a high-order thinking skills (HOTS)-based multiple-choice instrument to assess citizen participation in consumer education. Guided by Bannister and Monsma's framework, the instrument was developed through expert reviews and tested empirically using the Rasch model. The content validity, assessed by Aiken's V, CVR, and I-CVI, showed high expert agreement. Empirical validation involving 250 university students indicated excellent item reliability (0.99), acceptable person reliability (0.79), and appropriate item fit. All 24 items analyzed fell within acceptable MNSQ ranges and showed positive point-measure correlations. While item difficulties ranged from –2.23 to +1.87 logits, most were easier for the sample, suggesting potential refinement for high-ability students. The findings support the instrument's validity and reliability for assessing Higher Order Thinking Skills (HOTS) in citizen participation, providing a robust tool for assessing transformative consumer education.

Address for Correspondence: [1]putudiaharikusumadewi@uny.ac.id

Keywords: Consumer Education; Citizen Participation; Rasch Model; Content Validity

INTRODUCTION

Consumer education is an interdisciplinary field that aims to equip individuals with the knowledge, skills, and attitudes needed to make wise, ethical, and responsible consumption decisions. In its development, consumer education has not only focused on financial literacy and household management but also includes the social and political dimensions of consumption. McGregor emphasized that consumer education plays a strategic role in forming citizen-consumers (McGregor, 2016). These individuals, in particular, realize that every act of consumption has social implications and contributes to a democratic market system. One of the main dimensions in consumer education is citizen participation, namely the active involvement of consumers in public issues that affect their consumption lives. McGregor identifies citizen participation and advocacy as one of the three pillars of consumer education (McGregor, 2018), proposed by Bannister and Monsma, alongside decision-making and resource management (Bannister, 1983). However, in practice, this dimension is often marginalized, both in the curriculum and in assessment strategies. International standards, such as those of the OECD, tend to emphasize managerial aspects, so critical and participatory approaches have not received adequate attention.

In Indonesia, the low level of consumer awareness of rights and obligations, as revealed by Simanjuntak et al. (Simanjuntak & Sudibjo, 2019), as well as the lack of critical awareness of market structures that are not always fair (Oshio & Ihensekhien, 2019), indicate the urgency to reorient consumer education towards a more participatory direction. Unfortunately, higher education still predominantly adopts a conventional approach, with an emphasis on

functional skills such as financial planning or wise purchasing (Stewart & Yap, 2020) rather than on developing students' capacities as reflective and advocate citizens.

Gaps also occur in the assessment aspect. Most consumer education assessment instruments still measure low-level thinking skills (Lower Order Thinking Skills/LOTS). Instruments that can measure Higher Order Thinking Skills (HOTS), such as analyzing (C4), evaluating (C5), and creating (C6), especially in the context of citizen participation, are still very limited. HOTS is a crucial pillar in 21st-century education (Anderson & Krathwohl, 2001) and is essential for assessing students' ability to evaluate policies, critique market practices, and develop socially relevant advocacy strategies. Previous studies have developed various instruments in the realm of consumer education, but most of them still focus on financial literacy (Mandell, 2008) or general education program assessment (McCall, 1973); (Remmele, 1981) without explicitly touching on the citizen participation domain or using an Item Response Theory (IRT)-based validation approach.

IRT, especially the Rasch Model, enables the objective analysis of item characteristics, including the estimation of item difficulty, reliability, and the mapping of participant ability distributions. Clement's study in St. Lucia demonstrated that without critical and empowering consumer education, vulnerable communities struggle to navigate market complexity (Clement, 2013). This reinforces the urgency of providing valid, reliable, and socially relevant assessment instruments. This study aims to develop and test the construct validity and reliability of a HOTS-based dichotomous multiple-choice instrument to measure citizen participation in consumer education. The instrument is based on the framework of Bannister and Monsma (Bannister, 1983). It has been empirically validated using the Rasch Model approach to produce a measuring instrument that aligns with the needs of reflective, transformative, and competency-based consumer education in the 21st century.

METHOD

This study applied a research and development (R&D) approach adapted from Borg and Gall (1983), which involves seven key stages: preliminary study, planning, initial product development, expert validation, limited testing, main field testing, and final product revision. This approach was chosen to develop and empirically validate a high-order thinking skill (HOTS)-based multiple-choice assessment instrument designed to measure citizen participation within the context of consumer education.

Participants in this study were 250 undergraduate students from the Fashion Education Program at Universitas Negeri Yogyakarta who had completed a Consumer Education course. A purposive sampling technique was employed to ensure that participants had sufficient exposure to relevant course content. The sample size met the minimum requirements for Rasch model analysis using Item Response Theory (IRT). The developed instrument consisted of 30 dichotomous multiple-choice items designed to assess higher-order thinking Skills (HOTS) levels C4 (analyze), C5 (evaluate), and C6 (create), as outlined in Bloom's revised taxonomy (Anderson & Krathwohl, 2001). Item construction was guided by the conceptual framework of Bannister and Monsma (1982), which categorizes citizen participation into two main subdimensions: consumer protection and consumer advocacy.

To ensure content validity, the expert review was conducted involving three specialists in consumer education, educational evaluation, and HOTS pedagogy. The experts assessed each item for content relevance, clarity of wording, and cognitive level. The validity indices used included Aiken's V, Content Validity Ratio (CVR), and Item-Level Content Validity Index (I-CVI). Items with an Aiken's V of ≥ 0.80 were retained for further testing. Data collection was conducted online using a secure, time-bound digital form. Respondents were given 60 minutes to complete the test under standardized supervision protocols. For data analysis, the Rasch model (1PL IRT) was employed using the R software. Analyses included item-person fit

diagnostics (infit and outfit MNSQ), estimation of item difficulty in logits, testing unidimensionality through Principal Component Analysis of Residuals (PCAR), calculation of item and person reliability indices, and test information function analysis to evaluate measurement precision across varying ability levels.

RESULTS AND DISCUSSION

1. Development of HOTS-Based Instruments

In response to the gaps in HOTS assessment in the citizen participation domain identified in the preliminary study, this section presents the results of developing indicators and the cognitive structure of the instrument. This instrument was developed to evaluate students' abilities in consumer education in a critical, reflective, and creative manner. Based on international literature reviews (McGregor, 2016); (Bannister, 1983) and in-depth interviews with consumer education experts, citizen participation indicators are categorized into two main dimensions: consumer protection and consumer advocacy, which are further expanded into five subdimensions, namely: consumer rights & responsibilities, consumer laws, consumer assistance, consumer assertiveness, and consumer organization. Each item indicator is designed to measure three domains of higher-order thinking skills: analyzing (C4), evaluating (C5), and creating (C6), which correspond to the revised Bloom's Taxonomy (Anderson & Krathwohl, 2001). Table 1 shows an example of the integration between subdimensions, indicators, cognitive levels, and item numbers:

Table 1. Citizen Participation Dimensions

Dimensions	Sub Dimensions
Consumer Protection	Consumer Rights & Resposibilities
	Consumer Laws
	Consumer Assistance
Consumer Advocacy	Consumer Assertiveness
	Consumer Organization

To provide a systematic overview of the instruments developed, Table 2 shows the indicators that have been developed:

Table 2. Item Indicators and Sub Indicator

Sub Dimension	Item Indicator	Sub Indicator	Cognitive Level	Item Number
Consumer Rights & Responsibilities	Students analyze the rights and responsibilities of consumers, producers, and the government in protecting consumers.	Students analyze the differences in the rights and responsibilities of consumers, producers, and the government in terms of consumer protection.	C4	1
		Students connect the rights and responsibilities between parties in protecting consumers.	C4	2
		Students analyze the factors that most influence consumer protection.	C4	3
	Students evaluate the successes and failures of consumers, producers, and governments in fulfilling their responsibilities.	Students examine the fulfillment of each party's responsibilities in creating a good market for consumers.	C5	4
		Students criticized the imbalance of responsibility between the parties.	C5	5
	Students develop consumer protection efforts.	Students develop reflective efforts to protect consumers	C6	6
Consumer Laws	Students analyze legal steps in consumer loss cases	Students analyze the differences between online and conventional fraud	C4	7
		Students compile steps taken by consumers who are victims of online fraud.	C4	8
		Students analyze the implications of legal steps taken to protect consumers.	C4	9
	Students evaluate the impact of consumer protection laws	Students examine factors that influence the successful implementation of consumer protection laws.	C5	10

Sub Dimension	Item Indicator	Sub Indicator	Cognitive Level	Item Number
		Students evaluate the advantages and disadvantages of consumer protection laws.	C5	11
	Students develop improvement strategies for consumer protection	Students develop improvement strategies for consumer protection	C6	12
Consumer Assistance	Students analyze the role and function of consumer protection agents	Students analyze the differences between consumer protection agents and consumer dispute resolution institutions.	C4	13
		Students compile steps taken by consumer protection agents in handling consumer complaints.	C4	14
		Students analyze the success of consumer protection agencies in handling consumer losses.	C4	15
	Students evaluate the influence of consumer protection agencies	Students assess the effectiveness of consumer protection agencies in enhancing consumer awareness.	C5	16
		Students criticize shortcomings in consumer protection policies	C5	17
	Students develop the roles and functions of consumer protection agents	Students develop the roles and functions of consumer protection agents	C6	18
Consumer Assertiveness	Students analyze the steps of consumer action strategies	Students analyze the differences between informal and formal approaches to solving consumer problems.	C4	19
		Students organize the stages of making consumer complaints	C4	20
		Students analyze problem-solving steps.	C4	21
	Students evaluate the process of filing complaints or resolving consumer problems through legal processes.	Students examine obstacles to resolving consumer complaint cases.	C5	22
		Students identify Action strategies available to consumers to combat unfair business practices.	C5	23
	Students create consumer problem-solving strategies.	Students formulate strategies for resolving consumer problems, including how to file complaints and obtain compensation.	C6	24
Consumer Organization	Students analyze the role of consumer organizations in protecting consumer interests.	Students distinguish the roles played by consumer organizations in representing consumer interests.	C4	25
		Students organize information (goals and methods) about various consumer organizations.	C4	26
		Students attribute the impact of group actions, advocacy campaigns, or other collective efforts undertaken by consumer organizations towards protecting consumer rights.	C4	27
	Students evaluate the role of consumer organizations in representing consumer interests through group action, advocacy campaigns, or other collective efforts.	Students criticize the strategies taken by consumer organizations in achieving their goals.	C5	28
		Students evaluate the strengths and weaknesses of group actions, advocacy campaigns, or collective efforts undertaken by consumer organizations.	C5	29
	Students create new strategies to represent consumer interests through group action, advocacy campaigns, or other collaborative efforts.	Students develop innovative concepts to increase consumer participation in group action or advocacy campaigns.	C6	30

This instrument reflects a systematic mapping between cognitive aspects and the substance of contextual consumer education. For example, item C4 in consumer laws assesses students' ability to distinguish the legal context in fraud cases. In contrast, item C6 in consumer organizations encourages students to design advocacy campaign strategies based on social reality.

2. Content Validity Test

After the indicators are compiled based on a systematic conceptual framework, the next step is to ensure the validity of the instrument content. The content validity test aims to assess the extent to which the assessment items accurately represent the concept of citizen participation being measured in the context of consumer education. The validation process was conducted by three experts with expertise in consumer education, learning evaluation, and HOTS-based pedagogy. Each item is assessed based on three main aspects: content suitability, editorial clarity, and cognitive level. This assessment is analyzed quantitatively using three approaches, namely Aiken's V, Content Validity Ratio (CVR), and Item Content Validity Index (I-CVI).

Table 3. Content Validity Results

Item Number	V	CVR	CVI
1	0,9	1	1
2	0,75	1	1
3	0,9	1	1
4	1	1	1
5	0,65	1	0,8
6	0,6	1	1
7	0,65	1	0,8
8	0,85	1	1
9	0,7	1	0,8
10	0,85	1	1
11	0,55	1	1
12	0,75	1	1
13	0,85	1	1
14	0,65	1	0,6
15	0,9	1	1
16	0,7	1	1
17	0,9	1	1
18	0,95	1	1
19	0,75	1	1
20	0,65	1	0,6
21	0,65	1	1
22	0,85	1	1
23	0,8	1	1
24	0,75	1	1
25	0,65	1	0,6
26	0,55	1	1
27	0,95	1	1
28	0,85	1	1
29	1	1	1
30	0,85	1	1

As shown in Table 3, the test results show that out of 30 instrument items, the Aiken's V value ranges from 0.55 to 1.00, with an average of 0.78. The CVR value of the 30 items is 1.00, indicating complete consensus among experts regarding the essentiality of each item. The I-CVI analysis showed that 24 out of 30 items scored 1.00, indicating complete consensus on the relevance of the item. The other six items obtained I-CVI values between 0.60 – 0.80. Based on validation standards (Polit et al., 2007), items with I-CVI values <0.78 on a panel of

three people were not proceeded to the field trial stage. This finding confirms that most of the items have met substantial criteria in terms of content validity. High content validity also indicates that the instrument has a strong representation of the construct and good relevance to learning objectives. This strengthens the argument that the assessment developed is not only theoretically feasible but also practically credible for use in the context of higher education. Thus, the content validity test serves as an initial filter, ensuring that only truly qualified items are tested further in empirical validation using the Rasch Model.

3. Construct Validity and Reliability Analysis with the Rasch Model

After ensuring content validity, testing continued to the empirical validation stage to assess the construct validity and reliability of the instrument. A total of 250 students from the Fashion Education Study Program who had taken the Consumer Education course were involved as respondents. The data from the answers were analyzed using the Item Response Theory (IRT) approach, specifically the Rasch model, through the R software. Table 4 presents summary statistics for the suitability of items and persons to the Rasch model.

Table 4. Summary of Statistics

Statistical Measurement	Item	Person
Mean Measure (logit)	0.00	0.72
Reliability	0.99	0.79
Infit Mean Square	1.00	1.01
Outfit Mean Square	1.02	1.02
Mean Infit Z-Standard	0.05	-0.13
Mean Outfit Z-Standard	0.09	-0.18
Separation Index	9.00	2.30

The item reliability index of 0.99 and person reliability of 0.79 were achieved. The mean squared infit and outfit values for items and persons were within the recommended range of 0.5 to 1.5, confirming good fit between observed responses and model expectations. The separation index of 9.00 for items indicates the instrument can distinguish at least nine strata of item difficulty. The person separation index of 2.30 indicates the ability to classify respondents into approximately three levels of ability. All items yielded positive PTMEA correlation values (0.15 to 0.31). The range of item difficulty levels was from -2.23 to +1.87 logits. Table 5 presents a summary of Rasch diagnostics for each item.

Table 5. Summary of Rasch Diagnostics per Item

Item Number	Measure (logit)	Infit MNSQ	Outfit MNSQ	PTMEA CORR
1	-2.233	1.03	0.99	0.26
2	-1.630	0.94	0.98	0.29
3	-1.688	1.04	1.04	0.27
4	-1.260	1.05	1.07	0.24
5	-1.602	0.98	1.00	0.28
6	-1.003	0.98	1.02	0.30
7	-0.893	0.93	0.91	0.31
8	-0.702	1.02	1.10	0.26
9	-0.559	0.96	0.94	0.28
10	-0.765	0.97	0.97	0.29
11	-0.499	0.96	0.93	0.30
12	-0.187	1.03	1.04	0.22
13	-0.110	1.00	0.99	0.21
14	0.311	1.00	1.00	0.20
15	0.668	1.06	1.13	0.19
16	0.750	0.97	0.95	0.18
17	1.098	0.95	0.92	0.17
18	0.856	1.01	1.00	0.19
19	0.986	1.02	1.07	0.18
20	1.312	0.98	0.90	0.17

Item Number	Measure (logit)	Infit MNSQ	Outfit MNSQ	PTMEA CORR
21	1.546	1.04	1.10	0.16
22	1.808	1.04	0.97	0.15
23	1.872	1.04	1.16	0.15
24	1.747	0.99	0.98	0.15

Figure 1 presents the Wright Map (person-item map), which depicts the distribution of participants and items on a logit scale. Most respondents are in the mid to high ability range, while items are more dominantly distributed in the lower logit range.

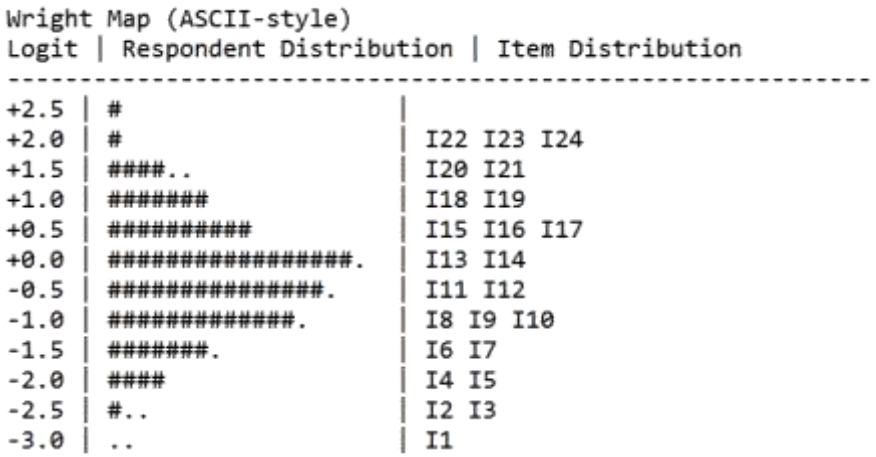

Figure 1. Wright Map Result

The Wright Map (Figure 1) shows that most respondents are in the mid to high ability range, while items are more dominantly distributed in the lower logit range. Principal Component Analysis of Residuals (PCAR) produced an eigenvalue of 1.62 in the first component, which is below the threshold of 2.0, indicating that the unidimensionality assumption is met.

Discussion

1. Development of HOTS-Based Instruments

The systematic mapping between cognitive aspects and the substance of contextual consumer education represents a significant advancement in assessment development. For example, item C4 in consumer laws assesses students' ability to distinguish the legal context in fraud cases, while item C6 in consumer organizations encourages students to design advocacy campaign strategies based on social reality.

This finding strengthens the argument that in-depth and theory-based indicator mapping not only produces conceptually valid assessments but also enables educators to evaluate more complex and strategically valuable competency dimensions. This aligns with McGregor's notion that consumer education should facilitate the transformation of students into active and reflective consumer citizens (McGregor, 2018). The indicator development stage plays a crucial role in building the foundation of an instrument that is not only relevant in content but also sensitive to the needs of 21st-century education in forming critical and empowered citizen-consumers.

2. Content Validity Test Results

The content validity results demonstrate substantial expert agreement on the instrument's quality. These values align with or exceed those reported in similar studies, such as Nurjanah et al., who achieved an Aiken's V value of ≥ 0.70 (Nurjanah et al., 2023), and

Anggraini et al., who reported a content validity coefficient of more than 0.87 for most items (Anggraini et al., 2020). Strong agreement among experts indicates that the instrument effectively captures the most important aspects of the assessment. High content validity also indicates that the instrument has a strong representation of the construct and good relevance to learning objectives. This strengthens the argument that the assessment developed is not only theoretically feasible but also practically credible for use in higher education contexts. The content validity test serves as an initial filter, ensuring that only truly qualified items are tested further in empirical validation using the Rasch Model.

3. Construct Validity and Reliability Analysis With The Rasch Model

The findings from the Rasch analysis provide strong empirical support for the construct validity and reliability of the developed instrument. The item reliability index of 0.99 exceeds the minimum threshold of 0.80 commonly accepted in psychometric research (Medvedev & Krägeloh, 2022). This high level of item reliability is consistent with previous instrument validation studies such as that conducted by Almeida et al., who reported an item reliability of 0.97 in a Rasch-based evaluation of a citizenship competency instrument for higher education students (Almeida et al., 2021).

The person reliability value of 0.79, while not as high as item reliability, is still in line with benchmarks reported by Ridzuan et al. (Ridzuan et al., 2020) and Adi et al. (Adi et al., 2022), who noted that individual reliability indices between 0.70 and 0.80 are sufficient for group-level decisions in educational assessment.

The fit statistics support the unidimensionality assumption of the Rasch model, aligning with studies by Rentiana et al. (Rentiana et al., 2024) and McLaughlin et al. (McLaughlin et al., 2023), which reported MNSQ infit and outfit values close to 1 as indication of strong item-model congruence in education-based constructs.

The separation indices are comparable to those reported by Sayin in validating an instrument assessing Turkish teacher candidate competency levels (Sayın & Şata, 2022). The positive PTMEA correlation values align with results from studies by Rosli et al. (Rosli et al., 2020) and Hazirah (Hazirah et al., 2021), highlighting the importance of positive PTMEA values as indicators of effective measurement direction.

The Wright Map results indicate that most items fall within the lower logit range compared to the distribution of respondents' abilities, which are concentrated at medium to high levels. This finding indicates that most items are relatively easy for the majority of respondents. This finding is consistent with those of Shakeel (Shakeel et al., 2023) and Yudha (Yudha, 2023), which show that item distributions that are too easy can limit the diagnostic function for students with high abilities.

The implications suggest the need for item revision or enrichment, particularly at the C6 (create) level, to enhance the instrument's ability to differentiate individuals with higher cognitive abilities. The addition of items with logit difficulty levels greater than 2.0 has been suggested in Triyani's study, where difficult items are needed to overcome the ceiling effect and increase discrimination in high-achieving groups (Triyani et al., 2021).

The information function analysis shows that the peak of measurement information is in the logit range of -1 to +1, indicating that the highest accuracy of this instrument is achieved with students of intermediate ability. This finding is consistent with results from Avinç, which shows that most Rasch-based measuring instruments have the highest effectiveness in measuring abilities in the leading population group but tend to be less precise at the distribution extremes (Avinç & Doğan, 2024). The PCAR results confirm unidimensionality, aligning with criteria proposed by Burcu (Burcu & Prof, 2023) and confirmed in previous validation studies by Park & Park (Park & Park, 2019) and Yudha (Yudha, 2023).

CONCLUSION

This study has successfully developed and tested the construct validity and reliability of a dichotomous multiple-choice instrument based on Higher Order Thinking Skills (HOTS) to measure citizen participation in the context of consumer education. The instrument is designed based on a theoretical framework and refers to the Revised Bloom's Taxonomy at levels C4 to C6. The results of content validation, as measured by Aiken's V, Content Validity Ratio (CVR), and Item-Level Content Validity Index (I-CVI), demonstrate a high level of agreement among experts, with most items achieving a V value greater than 0.80 and CVR and I-CVI reaching a value of 1.00. From the empirical validity test using the Rasch Model, all 24 items tested met the item fit criteria, with infit and outfit MNSQ values ranging from 0.90 to 1.13 and a positive correlation with the measured construct (PTMEA CORR = 0.15–0.31). Person and item reliabilities were 0.79 and 0.99, respectively, indicating a stable and sensitive instrument in differentiating students' abilities.

Despite these promising findings, this study has some limitations. The sample was limited to students from a single institution, which may have affected its generalizability. Future studies may consider testing the instrument across broader populations and integrating polytomous formats to gain a more profound insight. Nevertheless, this study provides a foundational tool to support reflective and transformative consumer education.

Further research is recommended to 1) Develop additional items with higher complexity; 2) Test the instrument on more diverse samples from various institutions or levels of education; and 3) Integrate quantitative and qualitative data (e.g., interviews or open-ended responses) to strengthen the interpretive validity of the instrument. Thus, this instrument is not only valid and reliable but also promises to make a significant contribution to enhancing the quality of HOTS-based assessments in transformative and contextual consumer education.

REFERENCE

Adi, N. R. M., Amaruddin, H., Adi, H. M. M., & A'yun, I. L. Q. (2022). Validity and Reliability Analysis Using the Rasch Model to Measure the Quality of Mathematics Test Items of Vocational High Schools. *Journal of Research and Educational Research Evaluation*, *11*(2), 103–113. https://doi.org/10.15294/JERE.V11I2.58835

Almeida, R. F., Pereira, N. D., Ribeiro, L. P., Barreto, R. P. G., Haik, M. N., & Camargo, P. R. (2021). Is the Disabilities of the Arm, Shoulder and Hand (DASH) Questionnaire Adequate to Assess Individuals With Subacromial Pain Syndrome? Rasch Model and International Classification of Functioning, Disability and Health. *Physical Therapy*, *101*(5). https://doi.org/10.1093/PTJ/PZAB065

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing : a revision of Bloom's taxonomy of educational objectives : complete edition*. https://eduq.info/xmlui/handle/11515/18824

Anggraini, D., Khumaedi, M., & Widowati, T. (2020). Validity and Reliability Contents of Independence Assessment Instruments of Basic Beauty Students for Class X SMK. *Journal of Research and Educational Research Evaluation*, *9*(1), 40–46. https://doi.org/10.15294/JERE.V9I1.42558

Avinç, E., & Doğan, F. (2024). Digital literacy scale: Validity and reliability study with the rasch model. *Education and Information Technologies*, *29*(17), 22895–22941. https://doi.org/10.1007/S10639-024-12662-7/FIGURES/17

Bannister, R. (1983). A Classification of Concepts in Consumer Education. *Http://Dx.Doi.Org/10.1177/019263658306746703*, *67*(467), 10–15. https://doi.org/10.1177/019263658306746703

Burcu, E., & Prof, A. (2023). A Review of Measurement Tools Developed and Adapted Based on the Rasch Model. *İnsan ve Sosyal Bilimler Dergisi*, *6*(2), 249–275. https://doi.org/10.53048/JOHASS.1369336

Clement, P. (2013). Consumer Education in a Developing Country: The Use of a Stratified Random Sample Survey. *International Journal of Management*, *30*(2).

Hazirah, N., Dzin, M., & Lay, Y. F. (2021). Assessing the Validity and Reliability of Science Multiple Choice Test Using Rasch Dichotomous Measurement Model. *Journal of Baltic Science Education*, *20*(6), 927–941. https://doi.org/10.33225/jbse/21.20.927

Mandell, L. (2008). Financial Literacy of High School Students. *Handbook of Consumer Finance Research*, 163–183. https://doi.org/10.1007/978-0-387-75734-6_10

McCall, C. (1973). DEVELOPMENT OF A TEST FOR ASSESSING COMPETENCY IN CONSUMER EDUCATION. In *McCall, C. L. (1973). DEVELOPMENT OF A TEST FOR ASSESSING COMPETENCY IN CONSUMER EDUCATION. The Pennsylvania State University.* https://search.proquest.com/openview/76e6bcb1563a1040c5c8c92a85501345/1?pq-origsite=gscholar&cbl=18750&diss=y

McGregor, S. L. T. (2016). Framing consumer education conceptual innovations as consumer activism. *International Journal of Consumer Studies*, *40*(1), 35–47. https://doi.org/10.1111/IJCS.12208

McGregor, S. L. T. (2018). Status of Consumer Education and Financial Education in Canada (2016). *Canadian Journal of Education*, *41*(2), 601–632. www.cje-rce.ca

McLaughlin, J. E., Angelo, T. A., & White, P. J. (2023). Validating criteria for identifying core concepts using many-facet rasch measurement. *Frontiers in Education*, *8*, 1150781. https://doi.org/10.3389/FEDUC.2023.1150781/BIBTEX

Medvedev, O. N., & Krägeloh, C. U. (2022). Rasch Measurement Model. *Handbook of Assessment in Mindfulness Research*, 1–18. https://doi.org/10.1007/978-3-030-77644-2_4-1

Nurjanah, S., Istiyono, E., Widihastuti, W., Iqbal, M., & Kamal, S. (2023). The Application of Aiken's V Method for Evaluating the Content Validity of Instruments that Measure the Implementation of Formative Assessments. *Journal of Research and Educational Research Evaluation*, *12*(2), 125–133. https://doi.org/10.15294/JERE.V12I2.76451

Oshio, L. E., & Ihensekhien, I. (2019). Effect of consumer education on consumer skills of adolescents in senior secondary schools in Nigeria: Implication for national peace and security. *International Journal of Educational Policy Research and Review*, *6*(2), 39–45. https://doi.org/10.15739/IJEPRR.19.006

Park, J. Y., & Park, E. Y. (2019). The Rasch Analysis of Rosenberg Self-Esteem Scale in Individuals With Intellectual Disabilities. *Frontiers in Psychology*, *10*, 462939. https://doi.org/10.3389/FPSYG.2019.01992/BIBTEX

Remmele, A. (1981). *The development and pilot testing of a model for assessing the effectiveness of innovative consumer education programs*. Oregon State University. https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/4b29b832x?locale=en

Rentiana, L. H., Dhini, U. R., Yuliawati, L., & Wulandari, E. T. (2024). Item Fit Analysis for Evaluating Academic Writing Performance With Rasch Measurement. *Attractive : Innovative Education Journal*, *6*(1), 645–653. https://doi.org/10.51278/AJ.V6I1.1144

Ridzuan, M. F., Lim, H. L., Ahmad Fozee, F. A., & Mohd Nasser, S. N. A. (2020). Rasch Analysis Model: Reliability and Validity of Superitem Test Instrument. *International Journal of Academic Research in Progressive Education and Development*, *9*(4). https://doi.org/10.6007/IJARPED/V9-I4/8166

Rosli, R., Abdullah, M., Choiro Siregar, N., Shazana, N., Hamid, A., Abdullah, S., Kok Beng, G., Halim, L., Mat Daud, N., Aminah Bahari, S., Majid, R. A., & Bais, B. (2020). Student

Awareness of Space Science: Rasch Model Analysis for Validity and Reliability. *World Journal of Education*, *10*(3), 170–177. https://doi.org/10.5430/wje.v10n3p170

Sayın, A., & Şata, M. (2022). Using Rasch analysis to examine raters' expertise Turkish teacher candidates' competency levels in writing different types of test items. *International Journal of Assessment Tools in Education*, *9*(4), 998–1012. https://doi.org/10.21449/IJATE.1058300

Shakeel, S. I., Al Mamun, M. A., & Haolader, M. F. A. (2023). Instructional design with ADDIE and rapid prototyping for blended learning: validation and its acceptance in the context of TVET Bangladesh. *Education and Information Technologies*, *28*(6), 7601–7630. https://doi.org/10.1007/S10639-022-11471-0/METRICS

Simanjuntak, M. F., & Sudibjo, N. (2019). MENINGKATKAN KETERAMPILAN BERPIKIR KRITIS DAN KEMAMPUAN MEMECAHKAN MASALAH SISWA MELALUI PEMBELAJARAN BERBASIS MASALAH [IMPROVING STUDENTS' CRITICAL THINKING SKILLS AND PROBLEM SOLVING ABILITIES THROUGH PROBLEM-BASED LEARNING]. *JOHME: Journal of Holistic Mathematics Education*, *2*(2), 108–118. https://doi.org/10.19166/JOHME.V2I2.1331

Stewart, C. R., & Yap, S. F. (2020). Low literacy, policy and consumer vulnerability: Are we really doing enough? *International Journal of Consumer Studies*, *44*(4), 343–352. https://doi.org/10.1111/IJCS.12569

Triyani, T., Mustika, M., Sakman, S., & Mantir, H. B. (2021). Development of High Order Thinking Skills Based Assessment Tools in Pancasila Course: Critical Digital Citizenship Oriented. *Jurnal Lingua Idea*, *12*(2), 136–144. https://doi.org/10.20884/1.JLI.2021.12.2.4942

Yudha, R. P. (2023). Higher Order Thinking Skills (HOTS) Test Instrument: Validity and Reliability Analysis With The Rasch Model. *EduMa: Mathematics Education Learning and Teaching*, *12*(1), 21–38. https://doi.org/10.24235/EDUMA.V12I1.9468