Volume 4, Issue 3 (2026) 374-388

EQUATOR SCIENCE JOURNAL (ESJ)
e-ISSN: 2988-3881
Journal homepage:
https://ejournal.edukhatulistiwa.com/index.php/ESJ

Development and Rasch Validation of an Energy Literacy and Sustainable Behavior Instrument for High School Students

Neng Sriatun[1], Didit Ardianto[1*], Indarini Dwi Pursitasari[1]
[1]Sekolah Pascasarjana, Universitas Pakuan, Indonesia

Received: September 14, 2026 | Revised: October 18, 2026 | Accepted: October 24, 2026
First Available Online: October 26, 2026 | Publication Date: October 31, 2026
DOI: https://doi.org/10.61142/esj.v4i3.502

*Corresponding author: diditardianto@unpak.ac.id

Keywords:
Education for sustainable development;
Energy literacy;
Instrument validation;
Rasch measurement;
Sustainable behavior

ABSTRACT

This study aimed to develop an instrument and obtain preliminary validity and psychometric evidence for measuring energy literacy and sustainable behavior among senior high school students. The instrument was developed using the ADDIE model and comprised three dimensions: energy knowledge, energy attitudes, and sustainable behavior. The initial instrument contained 64 items: 28 dichotomous knowledge items, 16 four-point Likert attitude items, and 20 four-point frequency behavior items. Four experts assessed content validity using the Content Validity Index (CVI), and Rasch analysis examined responses from 94 students (37 Grade X and 57 Grade XI). The I-CVI for all items and the S-CVI/Ave were both 1.000. Rasch analysis identified three misfitting items—Items 2 and 10 in energy knowledge and Item 9 in sustainable behavior. Removing these items resulted in a 61-item instrument. Person reliability ranged from 0.79 to 0.84, while item reliability ranged from 0.84 to 0.93. Wright maps indicated reasonable central item coverage but weaker targeting at higher levels of energy attitudes and sustainable behavior. Principal Component Analysis of Rasch Residuals indicated that the residual structure requires further investigation; strict unidimensionality was not conclusively established. Exploratory Welch's t-tests showed higher person measures among Grade XI students in energy knowledge ($p < .001$), energy attitudes ($p = .004$), and sustainable behavior ($p = .044$). Overall, the findings provide preliminary evidence rather than definitive validation. Larger and more diverse samples are needed to further examine the instrument's internal structure and generalizability.

This is an open access article under the CC BY-SA license.

INTRODUCTION

Energy is an important issue in sustainable development because modern life depends heavily on the availability and use of energy. Increasing energy demand, continued dependence on fossil fuels, and the environmental impacts of energy use require individuals not only to understand energy concepts but also to make responsible decisions and engage in sustainable practices. In this context, education plays a strategic role in preparing young people to understand energy issues as scientific, social, environmental, and sustainability-related challenges. Education for Sustainable Development (ESD) provides a framework that emphasizes the development of knowledge, skills, attitudes, values, and the capacity to take action toward sustainability (Agbedahin, 2019; Oe et al., 2022). Therefore, energy education should extend beyond conceptual understanding to include awareness, decision-making, and actions that support sustainable energy use.

One important outcome of energy education is energy literacy. Energy literacy is a multidimensional construct involving understanding of energy concepts and systems as well as affective and behavioral dimensions related to responsible energy use. DeWaters & Powers (2011, 2013) developed an energy literacy framework encompassing cognitive, affective, and behavioral domains. Their framework indicates that energy literacy cannot be reduced to knowledge alone because understanding energy issues needs to be accompanied by appropriate attitudes and behavioral tendencies. In the present study, this framework informed the development of the energy knowledge and energy attitudes dimensions. Energy knowledge addresses students' understanding of fundamental energy concepts, energy sources and their benefits, the importance of energy use, energy availability, societal and environmental impacts, and decision-making related to energy use. Energy attitudes focus on students' awareness and concern about energy issues, values toward energy use, and beliefs about their personal capability to respond to energy-related challenges. Previous research has likewise emphasized the multidimensional nature of energy literacy and the importance of assessing its different dimensions (Chodkowska-Miszczuk et al., 2021; Das & Richman, 2022). Recent research has expanded energy literacy studies to include learning strategies, modules, media, technology, and contextual approaches; however, strengthening students' energy literacy and developing appropriate measurement approaches remain important concerns, including in the Indonesian context (Dangkua et al., 2022; Sriatun et al., 2025; Usman et al., 2021).

An important challenge in energy education is that knowledge and positive attitudes do not necessarily translate into actual behavior. DeWaters & Powers (2011) demonstrated that the relationships among energy knowledge, affective dimensions, and energy-conservation behavior are not straightforward. This is consistent with research on pro-environmental behavior showing that knowledge, attitudes, intentions, and actual actions do not necessarily develop uniformly (Janmaimool & Khajohnmanee, 2019; Wang et al., 2021). The Theory of Planned Behavior (TPB) proposes that attitudes toward a behavior, subjective norms, and perceived behavioral control are associated with behavioral intentions and subsequent action (Ajzen, 1991). In environmental contexts, psychological and situational factors also influence whether behavioral intentions are translated into action (Conner & Norman, 2022; Jekauc et al., 2025). Recent evidence from Indonesia further highlights the relevance of literacy and social context to sustainable energy behavior. Hakam et al. (2025), for example, found that sustainability literacy and social support were positively associated with sustainable energy consumption behavior among university students. Safitri et al. (2026) also reported positive associations between local wisdom-based environmental education and pro-environmental behavior among high school students. These findings reinforce the importance of examining energy-related knowledge, attitudes, and reported behavior as related but distinct dimensions.

Although related, these concepts have different roles in explaining sustainable behavior. Attitudes refer to evaluations of a behavior, awareness concerns recognition of an issue, and perceived behavioral control concerns an individual's perceived ability to perform a behavior. Sustainable behavior, in contrast, concerns actions performed in everyday contexts. Accordingly, the present study treats energy knowledge, energy attitudes, and sustainable behavior as related but distinct dimensions. The sustainable behavior dimension includes energy-conscious actions and personal responsibility, as well as social and cultural norms, perceived behavioral control, the use of supporting facilities, and social action. The latter indicators represent contextual and psychosocial factors associated with

sustainable practices rather than direct measures of behavior alone. This operational distinction is intended to provide a broader assessment of students' energy-related knowledge, attitudes, and reported practices without assuming that these dimensions are interchangeable.

The need for such measurement becomes particularly relevant when energy education is developed within students' cultural and environmental contexts. Learning that connects scientific concepts with local social and cultural experiences can help students recognize the relevance of science to everyday life (King & Henderson, 2018; Zidny et al., 2020). Recent studies in Indonesian science and environmental education have examined the integration of local wisdom into sustainability-oriented learning. Safitri et al. (2026) found that local wisdom-based environmental education was positively associated with students' environmental attitudes and pro-environmental behavior. Similarly, Sudarmin et al. (2026) examined the integration of ethnoscience, socio-scientific issues, and virtual reality to support sustainable environmental literacy. These studies provide relevant educational contexts for incorporating local cultural knowledge into sustainability-oriented science learning. In the present study, the local wisdom of *Sampan Bidar* in Sambas is incorporated as a contextual element connecting physics, energy, culture, and sustainability. Sampan Bidar therefore provides a locally meaningful context for selected instrument items, connecting energy concepts with students' cultural and environmental experiences. Its inclusion represents a contextual feature of the instrument rather than an intervention whose educational effectiveness is evaluated in this study.

The development of such learning approaches requires an instrument capable of measuring relevant educational outcomes with adequate validity and reliability. Instrument development requires not only theoretical alignment but also evidence of content validity and empirical evidence regarding item functioning. Content validity can be examined through expert judgment, while empirical analysis can provide information about how items function in the target population (Boateng et al., 2018). The Content Validity Index (CVI) provides quantitative evidence regarding item relevance based on expert evaluation (Almanasreh et al., 2019; Polit et al., 2007). Furthermore, Rasch measurement provides information about item and person measures on a common logit scale and enables the evaluation of item fit, person and item reliability, separation indices, response-category functioning, dimensionality, and item-person targeting (Boone, 2016; Soeharto et al., 2024; Suryadi et al., 2025). Expert-based content validation and Rasch analysis provide complementary evidence: expert judgment evaluates the relevance of items to the intended constructs, whereas Rasch analysis examines empirical item functioning and measurement characteristics in the target sample. Although studies such as Sudarmin et al. (2026) have evaluated measurement properties in sustainability-oriented science education, their focus on learning media and sustainability literacy differs from the present study's focus on the development and preliminary Rasch validation of an energy-related instrument. Combining expert-based content validation with Rasch analysis can therefore provide complementary evidence for the quality of a newly developed instrument.

Recent studies have examined sustainable energy consumption behavior among Indonesian youth (Hakam et al., 2025), the relationship between local wisdom-based environmental education and pro-environmental behavior (Safitri et al., 2026), and the development of culturally contextualized learning approaches for sustainable environmental literacy (Sudarmin et al., 2026). These studies provide relevant evidence on sustainability-related learning and behavior, but their primary emphases differ from the present study, which focuses on the development and preliminary Rasch validation of an instrument assessing energy knowledge, energy attitudes, and sustainable behavior. The measurement challenge addressed in this study is to combine these three dimensions within one instrument while maintaining their conceptual distinction and examining how the items function empirically. The contribution of this study lies in the combination of a three-dimensional assessment framework, expert-based content validation, Rasch measurement, and selected locally contextualized items incorporating Sampan Bidar in Sambas, rather than in proposing a new theoretical framework. This study therefore aimed to develop and initially validate an instrument comprising three dimensions: energy knowledge, energy attitudes, and sustainable behavior. Energy knowledge and energy attitudes were grounded in the cognitive and affective dimensions of the framework proposed by DeWaters & Powers (2011, 2013), whereas sustainable behavior was developed as a related but distinct construct informed by the Theory of Planned Behavior and research on pro-environmental behavior. Content validity was evaluated by four experts using the CVI, while empirical validity and psychometric

characteristics were examined using Rasch measurement. Given that the empirical evaluation involved 94 students from a single senior high school, the findings are interpreted as preliminary evidence of content validity and measurement properties. Further validation with larger and more diverse samples is needed to examine the stability and generalizability of the instrument. The developed instrument is expected to provide an initial measurement tool for assessing students' energy literacy and sustainable behavior and to support further research on energy education integrating ESD and local wisdom.

METHOD

Research Design

This study employed a Research and Development (R&D) approach using the ADDIE model, consisting of Analysis, Design, Development, Implementation, and Evaluation (Dick et al., 2001; Wahira & Tolla, 2023). The model was used to develop and initially validate an instrument for measuring high school students' energy literacy and sustainable behavior. The development process included construct identification, instrument design, expert validation, empirical testing, and psychometric evaluation. The empirical evaluation was designed to provide preliminary evidence of item functioning and measurement properties, rather than definitive evidence of validity and generalizability across the broader senior high school student population.

Instrument Development Procedure

The instrument was developed through five ADDIE stages. Analysis involved reviewing relevant literature and identifying the measurement needs, dimensions, and indicators of energy literacy and sustainable behavior. Design involved developing the conceptual framework, item blueprint, response formats, and item specifications. Development included item construction, expert content validation, CVI analysis, and revision. Implementation involved empirical testing with high school students. Evaluation involved Rasch analysis and item refinement based on the psychometric results. The development procedure is presented in Figure 1. The empirical evaluation in this study represents an initial psychometric assessment; further evaluation using larger and more diverse samples is required to examine the stability and generalizability of the instrument.

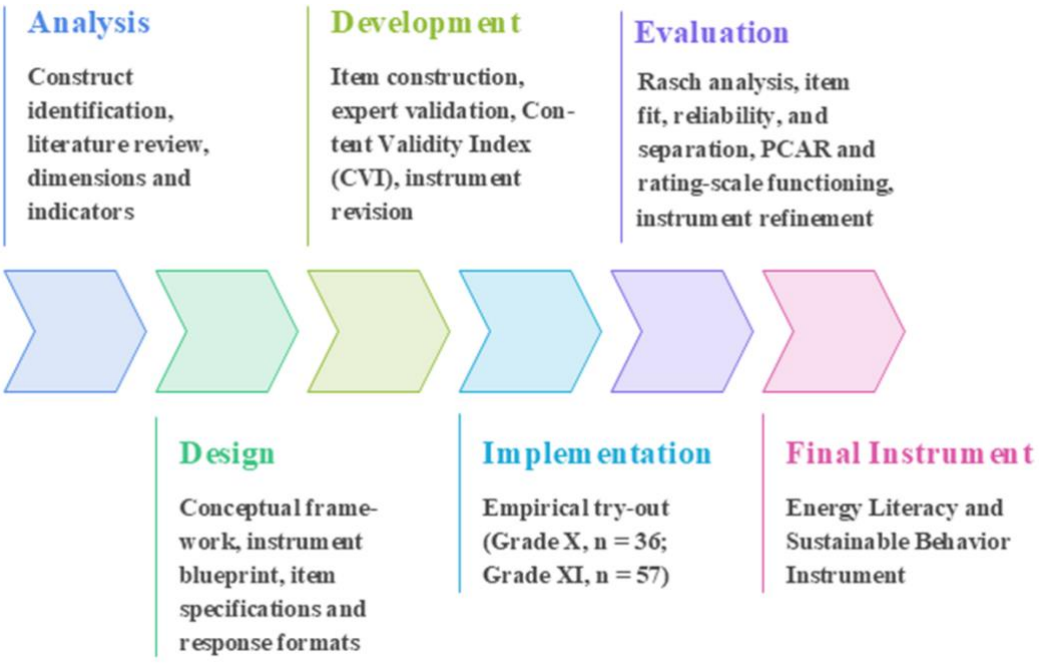


Fig.1. Instrument development and validation procedure based on the ADDIE model

The initial instrument consisted of 64 items comprising energy knowledge (28 dichotomous multiple-choice items), energy attitudes (16 four-point Likert items), and sustainable behavior (20 four-point frequency items). Energy knowledge and energy attitudes were grounded in the cognitive and affective dimensions of DeWaters & Powers (2011, 2013), respectively. Sustainable behavior was developed as a related but distinct construct informed by the Theory of Planned Behavior (Ajzen, 1991) and research on pro-environmental behavior. Selected items incorporated everyday and local contexts, including Sampan Bidar in Sambas. The initial instrument comprised three dimensions with distinct indicators and response formats, as summarized in Table 1.

Table 1. Summary of Instrument Dimensions and Representative Items

Dimension	Main aspects	Number of items	Response format	Representative item
Energy knowledge	Basic energy knowledge; energy sources and uses; importance of energy use; energy-use patterns and availability; impacts on society and the environment; energy-related decision making	28	Dichotomous multiple choice	"In the activity of paddling *Sampan Bidar*, a paddler lifts the paddle before moving it against the water to propel the boat. Based on the illustration, which statement best represents the law of conservation of mechanical energy?"
Energy attitudes	Awareness and concern about global energy issues; positive attitudes and values toward energy; self-belief in addressing energy issues	16	Four-point Likert	"Although it requires higher initial costs, the use of solar panels on water transportation such as boats is worth developing as a more sustainable energy solution."
Sustainable behavior	Energy-conscious behavior; personal responsibility; social and cultural norms; perceived behavioral control; access to supporting facilities; social action	20	Four-point frequency	"I use sunlight during the daytime before turning on lights at home or in the classroom."

The energy attitudes response categories were Strongly Agree, Agree, Disagree, and Strongly Disagree, whereas sustainable behavior used Always, Often, Rarely, and Never. Negative items were reverse-scored according to the direction of the construct.

Participants

The empirical validation involved 94 students from one senior high school in Sambas Regency, Indonesia, comprising 37 Grade X and 57 Grade XI students. A non-probability, class-based sampling approach was used. All students from the Grade XI science-track classes were included, while one Grade X class was selected based on the recommendation of the physics teacher, who also served as the class homeroom teacher. Grade X represented the target group for the intended use of the instrument, while Grade XI was included as a supplementary group to examine differences in person measures across adjacent grade levels. The sample was used for an initial examination of item functioning and the psychometric properties of the instrument in the intended educational context. Because the

participants were recruited from a single school using a non-random sampling approach, the findings are interpreted as preliminary evidence and should not be generalized to the broader senior high school student population without further validation. Subsequent studies involving larger samples from multiple schools and more diverse student populations are needed to examine the stability and generalizability of the instrument.

Content Validation

Content validity was assessed by four experts, comprising two physics education experts and two science literacy experts. The experts evaluated item relevance, clarity, and alignment with the intended constructs using a four-point rating scale. The Content Validity Index (CVI) was calculated using the item-level CVI (I-CVI) and scale-level CVI based on the average method (S-CVI/Ave). An I-CVI ≥ 0.78 and S-CVI/Ave ≥ 0.90 were considered acceptable (Azli Syam et al., 2026; Gao et al., 2025).

Data Analysis

Empirical data were analyzed using Rasch measurement. Energy knowledge was analyzed using the dichotomous Rasch model, whereas energy attitudes and sustainable behavior were analyzed using the Rating Scale Model (RSM). The analyses examined item fit, person and item reliability and separation, item measures, Principal Component Analysis of Rasch Residuals (PCAR), response-category functioning, and person-item targeting (Boone, 2016; X. Wang & Chen, 2024).

Item fit was evaluated using outfit MNSQ (0.50–1.50), ZSTD (−2.00–2.00), and PTMEA-CORR (0.40–0.85) (Boone, 2016; Tesio et al., 2024). For this study, an item was classified as misfitting only when all three indicators were simultaneously outside their specified ranges. This conjunctive rule was used as a study-specific decision criterion to identify items showing multiple indications of potential misfit, rather than as a universal Rasch standard. After item evaluation, person measures between Grade X and Grade XI were compared using the Welch t-test, with $p < 0.05$ indicating a statistically significant difference. These grade-level comparisons were exploratory and were not intended to provide evidence of item-level differential item functioning (DIF).

RESULTS AND DISCUSSION

Content Validation Results

Content validity was established through an expert evaluation involving four experts, comprising two experts in physics education and two experts in science literacy. The evaluation focused on the relevance, clarity, and appropriateness of each item in representing the constructs of energy knowledge, energy attitudes, and sustainable behavior. The initial instrument consisted of 64 items, including 28 items for energy knowledge, 16 items for energy attitudes, and 20 items for sustainable behavior.

All experts assigned scores of 3 or 4 to every item, with no ratings in categories 1 or 2. Based on the Content Validity Index (CVI), all items obtained an I-CVI of 1.000, while the overall S-CVI/Ave was also 1.000. These values exceeded the recommended criteria of I-CVI ≥ 0.78 and S-CVI/Ave ≥ 0.90 (Polit & Beck, 2006), indicating a high level of expert agreement regarding the relevance of the items. Similar findings have been reported in recent instrument validation studies, in which high I-CVI and S-CVI/Ave values were interpreted as evidence of strong expert agreement regarding item relevance and representativeness (Li et al., 2025; Ramezani & Mostafavi, 2025). Thus, the results provide strong initial evidence that the items adequately represented the intended content domains before empirical testing. The content validity results for the three dimensions of the instrument, based on expert evaluation, are presented in Table 2.

Table 2. Content Validity of the Instrument

Aspect	Initial Items	I-CVI Range	S-CVI/Ave	Interpretation
Energy Knowledge	28	1.000–1.000	1.000	Excellent
Energy Attitudes	16	1.000–1.000	1.000	Excellent

Aspect	Initial Items	I-CVI Range	S-CVI/Ave	Interpretation
Sustainable Behavior	20	1.000–1.000	1.000	Excellent
Overall	64	1.000–1.000	1.000	Excellent

The perfect agreement among the four experts indicates that no item was judged to have insufficient relevance at the content-validation stage. This finding supports proceeding to empirical validation because content validation provides an important basis for ensuring that the items are conceptually aligned with the intended construct domains (Boateng et al., 2018; Polit et al., 2007). However, high content validity does not by itself establish how items function empirically among the target respondents. Therefore, the instrument was subsequently examined using Rasch modeling to evaluate item functioning and other psychometric characteristics.

Rasch Validation Results

A total of 94 high school students participated in the empirical trial. Energy knowledge was analyzed using the dichotomous Rasch model, whereas energy attitudes and sustainable behavior were analyzed using the Rating Scale Model (RSM). The Rasch analysis examined item fit, person and item reliability, separation indices, dimensionality using Principal Component Analysis of Rasch residuals (PCAR), and response category functioning for the polytomous scales (Blackwood et al., 2023; Boone, 2016).

Item Fit and Instrument Refinement

Item fit was evaluated using outfit mean-square (MNSQ), outfit standardized residual (ZSTD), and PTMEA-CORR. The criteria were an outfit MNSQ of 0.50–1.50, outfit ZSTD of −2.00–2.00, and PTMEA-CORR of 0.40–0.85 (Boone, 2016; Saefi et al., 2020). An item was categorized as misfit when all three statistics simultaneously fell outside the specified ranges. The results of the initial Rasch analysis identified several items that did not meet the predetermined fit criteria, as presented in Table 3.

Table 3. Misfit Items in the Initial Rasch Analysis

Aspect	Item	Outfit MNSQ	Outfit ZSTD	PTMEA-CORR	Decision
Energy Knowledge	2	2.23	2.28	0.34	Removed
	10	1.82	2.14	0.33	Removed
Sustainable Behavior	9	2.01	5.71	0.07	Removed

Three items simultaneously exceeded the specified criteria across the three fit statistics. Item 2 and Item 10 in energy knowledge and Item 9 in sustainable behavior were therefore removed, followed by a re-analysis of the Rasch model. The removal of substantially misfitting items is consistent with previous Rasch-based studies in which item fit statistics were used to identify items that may not function consistently with the intended measurement construct (Soeharto, 2021; Soeharto & Csapó, 2022). After the re-analysis, no remaining item simultaneously met all three criteria for misfit. Consequently, the final instrument comprised 61 items: 26 energy knowledge items, 16 energy attitudes items, and 19 sustainable behavior items.

The refinement process demonstrates that content validity and empirical item functioning provide complementary forms of evidence. Although all 64 items received favorable expert judgments, empirical Rasch analysis identified three items that did not function consistently with the measurement model. This finding reinforces the importance of combining expert-based content evaluation with empirical item analysis when developing educational measurement instruments (Boateng et al., 2018).

Reliability, Separation, and Dimensionality

The re-analysis indicated that the final instrument demonstrated acceptable person and item reliability across the three dimensions. Energy knowledge showed person reliability of 0.84 and item reliability of 0.84, with person and item separation indices of 2.11 and 2.31, respectively. Energy attitudes demonstrated person reliability of 0.80 and item reliability of 0.92, with person separation of

2.03 and item separation of 3.31. Meanwhile, sustainable behavior showed person reliability of 0.79 and item reliability of 0.93, with person separation of 1.92 and item separation of 3.61 (Table 4).

Table 4. Summary of Rasch Analysis for the Final Instrument

Aspect	Items	Model	Person Reliability	Person Separation	Item Reliability	Item Separation
Energy Knowledge	26	Dichotomous Rasch	0.84	2.11	0.84	2.31
Energy Attitudes	16	RSM	0.80	2.03	0.92	3.31
Sustainable Behavior	19	RSM	0.79	1.92	0.93	3.61

Person separation exceeded 2.00 for energy knowledge and energy attitudes, suggesting that these dimensions could distinguish respondents across different levels of the measured traits. The lower person separation for sustainable behavior indicates comparatively more limited differentiation among respondents, despite its relatively high item reliability (0.93) and item separation (3.61). All item separation indices exceeded 2.00, suggesting that the item hierarchies were distinguishable in this sample. These findings are consistent with Rasch-based instrument studies in science education (Boone, 2016; Ismail et al., 2024). The implications of the lower person separation for sustainable behavior are considered alongside the person–item targeting findings below. Further refinement and testing with more heterogeneous samples may help improve the instrument's ability to distinguish respondents across the range of sustainable behavior.

The dimensionality of the final instrument was examined using Principal Component Analysis of Rasch Residuals (PCAR). The raw variance explained by measures was 30.2% for energy knowledge, 33.2% for energy attitudes, and 27.8% for sustainable behavior, while the first contrast values were 2.293, 3.253, and 2.770 eigenvalue units, respectively. Rasch studies commonly consider both explained variance and residual contrasts when examining dimensionality (Boone, 2016; Soeharto & Csapó, 2022). Although the Rasch measures explained approximately 28–33% of the variance, the first residual contrasts suggest that some systematic residual structure may remain unexplained. Therefore, these PCAR findings should be interpreted as preliminary evidence regarding the internal structure of the instrument rather than conclusive confirmation of unidimensionality.

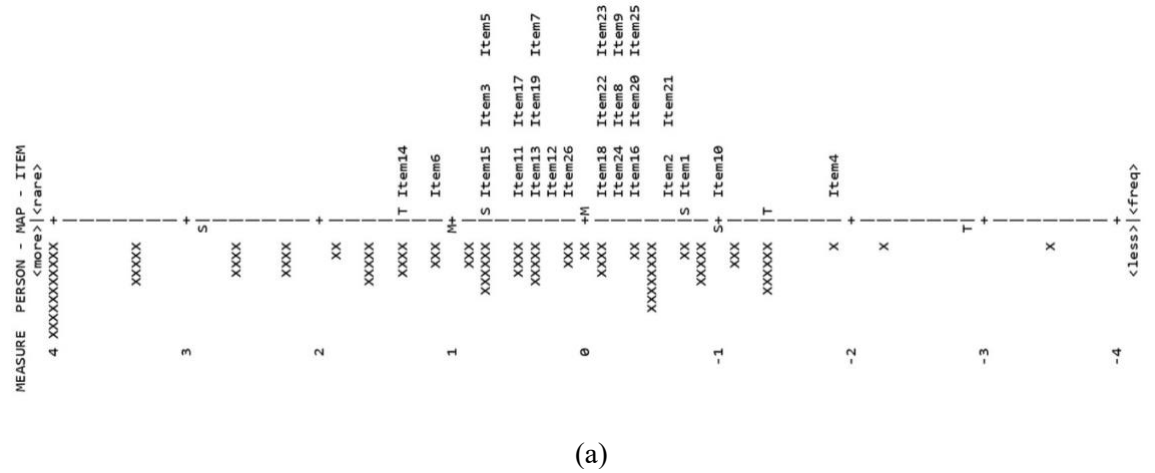
The highest first contrast was observed for energy attitudes (3.253), followed by sustainable behavior (2.770) and energy knowledge (2.293). The relatively higher residual structure in energy attitudes may reflect the breadth of this dimension, which includes awareness and concern, positive values toward sustainable energy, and beliefs about personal responsibility and capability. However, this explanation remains tentative because the first-contrast eigenvalue alone does not establish the substantive meaning of the residual structure. Thus, the PCAR results provide preliminary information about the intended measurement structure, while the internal structure of each dimension, particularly energy attitudes, remains to be examined further using additional residual diagnostics and larger, more diverse samples.

Wright Map

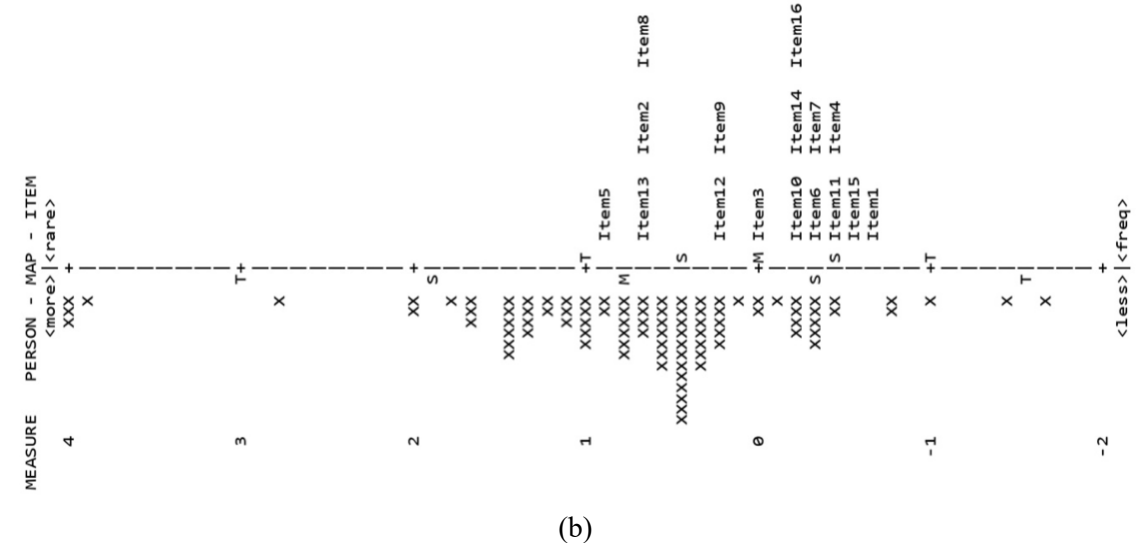
The Wright maps provide a visual comparison of student measures and item locations on the same logit scale, allowing the correspondence between respondent measures and item difficulty to be examined (Ismail et al., 2024; Isnawati et al., 2024; Soeharto & Csapó, 2022). As shown in Figure 2, energy knowledge showed relatively reasonable correspondence between person and item distributions in the central range, although several students were located beyond the upper and lower ends of the item distribution. This indicates adequate coverage for students in the central range but less coverage at the extremes.

For energy attitudes and sustainable behavior, the mean person measures were above the mean item measures, with several students located above the main range of item locations. This indicates under-targeting at the higher end of both constructs, suggesting that the current items provide less measurement coverage for students with higher levels of energy attitudes and sustainable behavior.

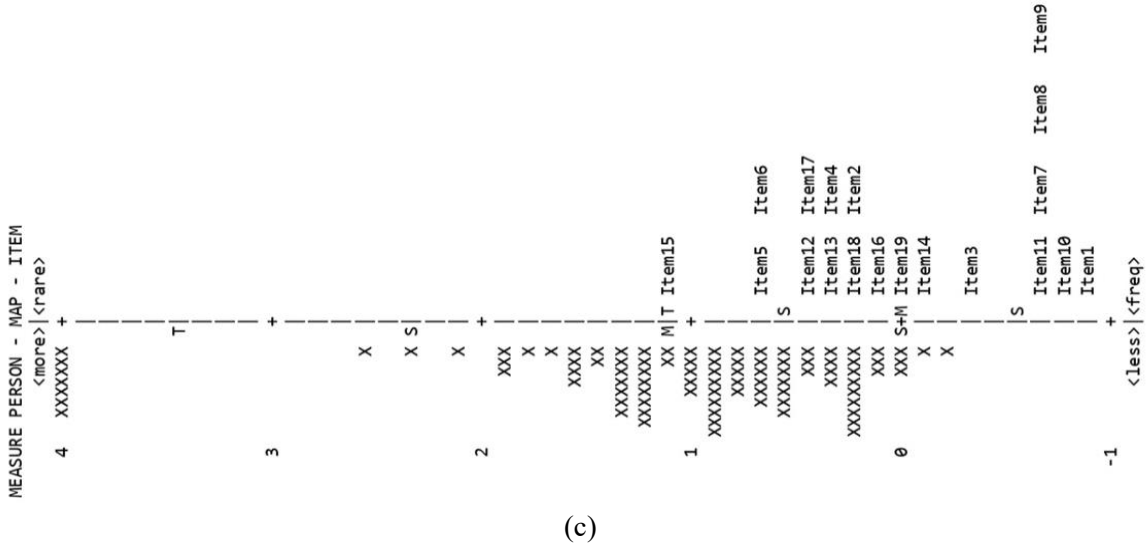
Previous Rasch studies have similarly used person-item maps to examine the correspondence between respondent measures and item difficulty and to identify areas for further item development (Isnawati et al., 2024; Soeharto et al., 2024).



(a)



(b)



(c)

Fig.2. Wright maps showing the distribution of student measures and item measures for (a) energy knowledge, (b) energy attitudes, and (c) sustainable behavior

Overall, the Wright maps indicate reasonable item coverage in the central range but less adequate targeting at the upper end, particularly for energy attitudes and sustainable behavior. This finding complements the reliability and separation results by identifying a specific area for further refinement. Future development should consider adding items representing higher levels of the constructs to extend the measurement range and improve targeting (Ismail et al., 2024; Isnawati et al., 2024).

Comparison of Person Measures by Grade Level

As an exploratory analysis, person measures were compared between Grade X and Grade XI students to examine whether the instrument could capture differences in the measured constructs across adjacent grade levels. Welch's t-test was used to compare the two independent groups because it does not require the assumption of equal population variances. The comparisons were interpreted as exploratory rather than confirmatory, and no adjustment for multiple testing was applied. The results are presented in Table 5.

Table 5. Comparison of Person Measures between Grade X and Grade XI Students

Aspect	Mean Difference	95% Confidence Interval	Welch's *t*	df	*p*	Cohen's *d*
Energy knowledge	1.531	[0.818, 2.245]	4.27	86.0	< .001	0.884
Energy attitudes	0.618	[0.204, 1.031]	2.97	89.7	.004	0.589
Sustainable behavior	0.426	[0.012, 0.840]	2.05	81.9	.044	0.400

Note: Mean differences are calculated as Grade XI minus Grade X**.** CI = confidence interval.

The results showed that Grade XI students had higher person measures than Grade X students across all three dimensions, with statistically significant differences in energy knowledge ($t = 4.27$, $df = 86.0$, $p < .001$), energy attitudes ($t = 2.97$, $df = 89.7$, $p = .004$), and sustainable behavior ($t = 2.05$, $df = 81.9$, $p = .044$). The corresponding Cohen's d values were 0.884, 0.589, and 0.400, respectively, indicating that the largest standardized difference was observed in energy knowledge, followed by energy attitudes and sustainable behavior**.** The 95% confidence intervals for the mean differences were [0.818, 2.245], [0.204, 1.031], and [0.012, 0.840], respectively. These estimates provide information about both the magnitude and uncertainty of the observed differences. However, the confidence interval for sustainable behavior approached zero, suggesting that the estimated difference in this dimension was comparatively small and less precisely estimated.

The larger difference in energy knowledge may reflect differences in students' exposure to science and energy-related learning experiences as they progress through secondary education. However, the smaller differences in energy attitudes and sustainable behavior indicate that the three dimensions may not develop at the same rate. This interpretation is consistent with the multidimensional nature of energy literacy, in which knowledge, attitudes, and behavior are related but do not necessarily develop uniformly (DeWaters & Powers, 2011; Sriatun et al., 2025). The relatively smaller difference in sustainable behavior may also reflect the influence of factors beyond knowledge, including subjective norms, perceived behavioral control, and contextual opportunities to act (Ajzen, 1991; Raghu & Rodrigues, 2022). These explanations are plausible interpretations rather than mechanisms directly tested in the present study.

Overall, the higher person measures among Grade XI students provide preliminary evidence of differences between the two sampled grade groups, with the largest observed difference in energy knowledge. However, these differences should not be interpreted as evidence that grade progression itself causes higher energy knowledge, energy attitudes, or sustainable behavior. They may also reflect differences in prior learning experiences, curriculum exposure, or school context. The findings should also be interpreted cautiously because the comparisons involved three statistical tests without a multiple-testing adjustment. In particular, the result for sustainable behavior ($p = .044$) was close to the conventional .05 threshold**.** In addition, the study involved only 94 students from a specific educational setting, with Grade XI serving as a supplementary group for examining measurement differences. Further studies with larger and more diverse samples are needed to evaluate the stability and generalizability of these findings. Future research should also determine whether the observed

differences remain after an appropriate multiple-testing correction and are replicated in independent samples.

Limitations

This study has several limitations. First, the empirical validation involved 94 students from a single school in Sambas Regency, Indonesia, recruited through non-probability sampling; therefore, the findings represent preliminary evidence and require further validation with larger, more diverse samples. Second, the instrument incorporated local elements, including Sampan Bidar Sambas, which may require adaptation for other cultural and educational contexts. Third, sustainable behavior was assessed through self-reports, which may be affected by social desirability and may not fully reflect actual behavior. Finally, two energy knowledge items (Items 2 and 10) and one sustainable behavior item (Item 9) were removed based on the study-specific misfit criteria. Their suitability and the internal structure of the instrument warrant further investigation in future studies.

CONCLUSION AND SUGGESTION

This study developed and initially validated an instrument for measuring high school students' energy literacy and sustainable behavior using a Research and Development (R&D) approach with the ADDIE model. The initial instrument comprised three dimensions: energy knowledge (28 dichotomous items), energy attitudes (16 polytomous items), and sustainable behavior (20 polytomous items). Content validation by four experts yielded an I-CVI of 1.000 for all items and an S-CVI/Ave of 1.000, indicating excellent content validity. Rasch analysis demonstrated satisfactory measurement quality across the three dimensions, with final person reliability ranging from 0.79 to 0.84 and item reliability ranging from 0.84 to 0.93. After item-fit evaluation, Items 2 and 10 in energy knowledge and Item 9 in sustainable behavior were removed, resulting in a final 61-item instrument. The dimensionality analysis provided preliminary evidence of the intended measurement structure, although the residual structure of energy attitudes warrants further investigation.

Comparison of person measures showed that Grade XI students had significantly higher measures than Grade X students across all three dimensions, including energy knowledge, energy attitudes, and sustainable behavior. The largest difference was observed in energy knowledge, followed by energy attitudes and sustainable behavior. These findings provide preliminary evidence of differences in person measures between the two sampled grade groups. However, the comparisons were exploratory and should not be interpreted as definitive evidence of the instrument's validity or as evidence that grade progression causes higher levels of these constructs.

Future studies should further examine the three removed items through substantive review and empirical testing before considering their potential reintegration. Validation with larger and more diverse samples is also needed to examine the stability and generalizability of the instrument. In addition, future research should improve item targeting, particularly at the higher levels of energy attitudes and sustainable behavior, and complement self-reported sustainable behavior with observational or other behavioral indicators. Further investigation of the internal structure of the three dimensions using larger samples and complementary validation approaches is also recommended.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS STATEMENT

Conceptualization, N.S., D.A., and I.D.P.; methodology, N.S., D.A., and I.D.P.; data analysis, N.S.; instrument development, N.S.; supervision, D.A. and I.D.P.; writing—original draft, N.S.; writing—review and editing, N.S., D.A., and I.D.P.; critical review and substantive input, D.A. and I.D.P.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Prof. Dr. Anna Permanasari, M.Si., Dr. Irvan Permana, M.Pd., and Dr. Nurhayati, M.Si., M.Pd. for their valuable feedback and suggestions as instrument validators during the development of this study. The authors also sincerely thank Jaka Afriana, M. Pd. for his assistance with data analysis.

DECLARATION OF GENERATIVE AI SOURCES

During the preparation of this work, the authors used ChatGPT to assist with language editing and manuscript preparation. The authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

Agbedahin, A. V. (2019). Sustainable development, education for sustainable development, and the 2030 Agenda for Sustainable Development: Emergence, efficacy, eminence, and future. *Sustainable Development, 27*(4), 669–680. https://doi.org/10.1002/sd.1931

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy, 15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Azli Syam, N. A., Isa, N. K. M., Yunos, M. Y. M., Isa, N. J. M., & Sormin, D. (2026). Content validity of a sustainable food management assessment instrument for island hotels using I-CVI and mean analysis. *Discover Applied Sciences, 8*(3), Article 331. https://doi.org/10.1007/s42452-026-08340-7

Blackwood, J., Carpentier, S., Deng, W., & Van de Winckel, A. (2023). Preliminary Rasch analysis of the multidimensional assessment of interoceptive awareness in adults with stroke. *PLOS ONE, 18*(6), Article e0286657. https://doi.org/10.1371/journal.pone.0286657

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

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

Chodkowska-Miszczuk, J., Kola-Bezka, M., Lewandowska, A., & Martinát, S. (2021). Local communities' energy literacy as a way to rural resilience—An insight from inner peripheries. *Energies, 14*(9), Article 2575. https://doi.org/10.3390/en14092575

Conner, M., & Norman, P. (2022). Understanding the intention-behavior gap: The role of intention strength. *Frontiers in Psychology, 13*, Article 923464. https://doi.org/10.3389/fpsyg.2022.923464

Dangkua, T., Mooduto, Y., & Tilome, A. (2022). Energy literacy education characteristics in Gorontalo City, Indonesia: Cognitive scale. *Journal La Lifesci, 3*(2), 82–91. https://doi.org/10.37899/journallalifesci.v3i2.608

Das, R. R., & Richman, R. (2022). The development and application of a public energy literacy instrument. *Canadian Journal of Science, Mathematics and Technology Education, 22*(1), 42–67. https://doi.org/10.1007/s42330-022-00196-4

DeWaters, J. E., & Powers, S. E. (2011). Energy literacy of secondary students in New York State (USA): A measure of knowledge, affect, and behavior. *Energy Policy, 39*(3), 1699–1710. https://doi.org/10.1016/j.enpol.2010.12.049

DeWaters, J., & Powers, S. (2013). Establishing measurement criteria for an energy literacy questionnaire. *The Journal of Environmental Education, 44*(1), 38–55. https://doi.org/10.1080/00958964.2012.711378

Dick, W., Carey, L., & Carey, J. O. (2001). *The systematic design of instruction* (5th ed.). Longman.

Gao, Y., Chen, X., Zhang, S., Yang, J., Liu, R., Moyle, W., Pu, L., & Shen, W. (2025). Confirmatory factor analyses of the Mandarin Chinese version of the perceived stressors in intensive care units among healthcare professionals. *Frontiers in Public Health, 13*, Article 1434440. https://doi.org/10.3389/fpubh.2025.1434440

Hakam, L. I., Hakam, D. F., Wiryono, S. K., & Rahadi, R. A. (2025). The impact of sustainability literacy, social support, and attitudes towards innovation on sustainable energy consumption among Indonesian youth. *Frontiers in Sustainable Cities, 7*, Article 1606031. https://doi.org/10.3389/frsc.2025.1606031

Ismail, I., Riandi, R., Kaniawati, I., Sopandi, W., Supriyadi, S., Suhendar, S., & Hidayat, F. A. (2024). Gender roles in understanding and implementing green energy technology in Indonesian schools: Rasch analysis. *Qubahan Academic Journal, 4*(3), 298–314. https://doi.org/10.48161/qaj.v4n3a752

Isnawati, I., Sriyati, S., Agustin, R. R., Supriyadi, S., Kasi, Y. F., & Ismail, I. (2024). Analysis of question difficulty levels based on science process skills indicators using the Rasch model. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah, 9*(1), 31–41. https://doi.org/10.24042/tadris.v9i1.22410

Janmaimool, P., & Khajohnmanee, S. (2019). Roles of environmental system knowledge in promoting university students' environmental attitudes and pro-environmental behaviors. *Sustainability, 11*(16), Article 4270. https://doi.org/10.3390/su11164270

Jekauc, D., Voelkle, M. C., Giurgiu, M., & Nigg, C. R. (2025). Unveiling the multidimensional nature of the intention–behavior gap: Dynamics of intentions and behaviors. *European Journal of Health Psychology, 32*(1), 34–50. https://doi.org/10.1027/2512-8442/a000162

King, D., & Henderson, S. (2018). Context-based learning in the middle years: Achieving resonance between the real-world field and environmental science concepts. *International Journal of Science Education, 40*(10), 1221–1238. https://doi.org/10.1080/09500693.2018.1470352

Li, C., Lin, Y., Guo, H. Y., Zhao, H., Xiong, J., Xiao, X. Y., & Zhang, L. (2025). Psychometric evaluation of the Online Education Student Satisfaction Scale: A translation and validation study among Chinese nursing students. *BMC Nursing, 24*(1), Article 1193. https://doi.org/10.1186/s12912-025-03864-6

Oe, H., Yamaoka, Y., & Ochiai, H. (2022). A qualitative assessment of community learning initiatives for environmental awareness and behaviour change: Applying UNESCO Education for Sustainable Development (ESD) framework. *International Journal of Environmental Research and Public Health, 19*(6), Article 3528. https://doi.org/10.3390/ijerph19063528

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459–467. https://doi.org/10.1002/nur.20199

Raghu, S. J., & Rodrigues, L. L. R. (2022). Solid waste management behavior among the student community: Integrating environmental knowledge and situational factors into the theories of

planned behavior and value belief norm. *Journal of Environmental Planning and Management, 65*(10), 1842–1874. https://doi.org/10.1080/09640568.2021.1949969

Ramezani, S. G., & Mostafavi, Z. S. (2025). Developing and validating a comprehensive scale for accreditation standards and quality assurance in e-learning institutions. *Education and Information Technologies, 30*(15), 21139–21187. https://doi.org/10.1007/s10639-025-13587-5

Saefi, M., Fauzi, A., Kristiana, E., Adi, W. C., Muchson, M., Setiawan, M. E., Islami, N. N., Ningrum, D. E. A. F., Ikhsan, M. A., & Ramadhani, M. (2020). Validating of knowledge, attitudes, and practices questionnaire for prevention of COVID-19 infections among undergraduate students: A Rasch and factor analysis. *Eurasia Journal of Mathematics, Science and Technology Education, 16*(12), Article em1926. https://doi.org/10.29333/ejmste/9352

Safitri, D., Ibrahim, N., Sujarwo, S., Marini, A., Pratama, I., Dewiyani, L., Julia, V., Saputro, R. H., Muharrani, N. P., Sudrajat, A., & Marfu, A. (2026). Understanding pro environmental behavior through local wisdom and AR based environmental education using PLS SEM and MGA approach. *Discover Sustainability, 7*(1), Article 633. https://doi.org/10.1007/s43621-026-02931-5

Soeharto, S. (2021). Development of a diagnostic assessment test to evaluate science misconceptions in terms of school grades: A Rasch measurement approach. *Journal of Turkish Science Education, 18*(3), 351–370. https://doi.org/10.36681/tused.2021.78

Soeharto, S., & Csapó, B. (2022). Assessing Indonesian student inductive reasoning: Rasch analysis. *Thinking Skills and Creativity, 46*, Article 101132. https://doi.org/10.1016/j.tsc.2022.101132

Soeharto, S., Martono, M., Hairida, H., Akhmetova, A., Arifiyanti, F., Csapó, B., & Charalambous, C. (2024). The metacognitive awareness of reading strategy among pre-service primary teachers and the possibility of rating improvement using Rasch analysis. *Studies in Educational Evaluation, 80*, Article 101319. https://doi.org/10.1016/j.stueduc.2023.101319

Sriatun, N., Ardianto, D., Permana, I., & Afriana, J. (2025). Strategies for promoting energy literacy in physics education: Insights from a systematic literature review. *Journal of Innovative Science Education, 14*(1), 54–63. https://doi.org/10.15294/jise.v14i1.20856

Sudarmin, S., Ariyatun, A., Permanasari, A., Hutagalung, F. D., Pujiastuti, R. S. E., Sari, D. S., & Asmaningrum, H. P. (2026). Evaluating the validity and effectiveness of the ethno-SSI project–integrated virtual reality in improving sustainable environmental literacy. *European Journal of STEM Education, 11*(1), Article 33. https://doi.org/10.20897/ejsteme/18731

Sukarelawan, M. I., Jumadi, Kuswanto, H., Soeharto, & Hikmah, F. N. (2021). Rasch analysis to evaluate the psychometric properties of junior metacognitive awareness inventory in the Indonesian context. *Jurnal Pendidikan IPA Indonesia, 10*(4), 486–495. https://doi.org/10.15294/jpii.v10i4.27114

Suryadi, N. R. S., Nurmegawati, L., Mu'aziyah, S. E. S., & Perdani, A. S. (2025). Rasch model for analysis of scientific attitude instruments in the context of secondary school science education. *Equator Science Journal*, *3*(2), 98–106. https://doi.org/10.61142/esj.v3i2.234

Tesio, L., Caronni, A., Simone, A., Kumbhare, D., & Scarano, S. (2024). Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment. *Disability and Rehabilitation, 46*(3), 604–617. https://doi.org/10.1080/09638288.2023.2169772

Usman, M., Suyanta, Pujianto, & Huda, K. (2021). Energy literacy of junior high school students in Indonesia: A preliminary study. In *Proceedings of the 6th International Seminar on Science Education (ISSE 2020)* (pp. 609–614). Atlantis Press. https://doi.org/10.2991/assehr.k.210326.088

Wahira, W., & Tolla, I. (2023). Analysis of the needs for developing the competence of elementary school supervisors through analysis design development implementation evaluation (ADDIE) model. *Kasetsart Journal of Social Sciences, 44*(4), 1303–1312. https://doi.org/10.34044/j.kjss.2023.44.4.34

Wang, H., Li, J., Mangmeechai, A., & Su, J. (2021). Linking perceived policy effectiveness and proenvironmental behavior: The influence of attitude, implementation intention, and knowledge. *International Journal of Environmental Research and Public Health, 18*(6), Article 2910. https://doi.org/10.3390/ijerph18062910

Wang, X., & Chen, S. (2024). Psychometric evaluation of the extended inclusion of nature in self scale in Chinese adults through the Rasch analysis. *People and Nature, 6*(6), 2431–2448. https://doi.org/10.1002/pan3.10726

Zidny, R., Sjöström, J., & Eilks, I. (2020). A multi-perspective reflection on how indigenous knowledge and related ideas can improve science education for sustainability. *Science & Education, 29*(1), 145–185. https://doi.org/10.1007/s11191-019-00100-x