15 (1) (2026) 1-12
Unnes Science Education Journal
Accredited Sinta 2
https://journal.unnes.ac.id/journals/usej

Developing a Minimum Competency Assessment in Science Literacy: A Rasch-Model Approach on Elements, Compounds, and Mixtures

Nirma Yuliana Saputri, Rizki Nor Amelia✉

DOI: http://dx.doi.org/10.15294/usej.v13i1.34242

Universitas Negeri Semarang, Indonesia

Article Info

Submitted 2025-09-24
Revised 2025-11-12
Accepted 2026-04-29

Keywords

Minimum competency assessment; Rasch model; Science literacy

Abstract

This study aims to develop a Minimum Competency Assessment (AKM) instrument for science literacy on the topic of elements, compounds, and mixtures using the Rasch Model. The research followed the 4D development model (Define, Design, Development, and Disseminate). The instrument initially consisted of 25 AKM-format items validated by five junior high school science teachers through expert appraisal using Aiken's V. The content validity results indicated that 22 items were valid, while three were eliminated for not aligning with the science literacy indicators. Empirical testing was conducted with 175 eighth-grade students from SMP Negeri 1 Juwana. The Rasch analysis showed high reliability (item reliability = .94; model reliability = .95), and all items met the model fit criteria based on the Infit and Outfit MNSQ values. The item difficulty ranged from very easy to very difficult (-1.28 to +1.60 logit), demonstrating a balanced distribution. The developed AKM-based science literacy instrument is valid, reliable, and pedagogically applicable. It can serve as both an evaluation tool and a learning support instrument that fosters critical thinking and scientific reasoning skills, aligning with the AKM and PISA 2025 frameworks.

How to Cite

Saputri, N. Y. & Amelia, R. N. (2026). Developing a Minimum Competency Assessment in Science Literacy: A Rasch-Model Approach on Elements, Compounds, and Mixtures. *Unnes Science Education Journal*, 15(1), 1-12.

✉ Correspondence Author:
E-mail: rizkinoramelia@mail.unnes.ac.id

p-ISSN 2252-6617
e-ISSN 2502-6232

INTRODUCTION

The transformation of education system in Indonesia emphasizes strengthening students' fundamental competencies through flexible and contextual learning approaches (Lubis et al., 2024; Mazidah & Suwarna, 2025). As part of its evaluative policy, the government has established the Minimum Competency Assessment (Asesmen Kompetensi Minimum or AKM) to measure students' proficiency in reading literacy, numeracy, and science literacy (Annisa et al., 2024; Ayuningtyas et al., 2023; Hawariah et al., 2024; Permatasari et al., 2023). The development of AKM based on Program for International Student Assessment (PISA) framework, objectively evaluates students' skills, particularly in science literacy (Hudson et al., 2021); Kim & Kim, 2021).

Science literacy represents the ability to think critically, solve problems, and make decisions based on scientific evidence relevant to daily life (Broderick, 2025; Valentini et al., 2022; Valladares, 2021). Following global developments, PISA 2025 updates the science literacy indicators by incorporating digital competencies and a deeper understanding of science (Agnesa et al., 2023; Deta et al., 2025; Pratama et al., 2024). These updates emphasize three key competencies: (1) explaining scientific phenomena, (2) evaluating and designing scientific investigations and interpreting data and evidence scientifically, and (3) using scientific information for decision-making and action (OECD, 2025).

However, Indonesian students' achievement in science literacy remains relatively low. The PISA 2023 score was 383, which is below the international average, indicating limitations in contextual understanding and the application of scientific concepts (OECD, 2023). This underachievement is influenced by students' limited reading habits and lack of practice with real-life-based questions (Ducha & Husna, 2025); Waini et al., 2024). Teachers have also not fully integrated science literacy indicators into assessment instruments, as questions still emphasize memorization rather than application of concepts to authentic problems (Faisal et al., 2023). Indonesia's low performance in science literacy underscores the urgent need for contextual and valid assessment instruments that align with AKM and the updated PISA framework.

Observations at SMP Negeri 1 Juwana indicate that teachers often use multiple-choice questions for practicality. Consequently, students are less familiar with AKM-based question formats, particularly those addressing abstract topics such as elements, compounds, and mixtures. These topics are rarely assessed through contextualized literacy questions, leading students to memorize symbols and definitions rather than understanding their real-world applications (Alwalya & Rendy, 2024). Furthermore, science learning outcomes require students to explain concepts, identify types, and analyze separation methods of mixtures, which are essential components of science literacy relevant to environmental issues (Eralita & Setiawan, 2022).

This condition suggests the science literacy measurement tools currently used are not yet aligned with the characteristics of AKM and the PISA 2025 indicators. The available instruments still rely on conventional formats make it difficult for students to engage in complex and contextual reasoning (Dwiwulandari et al., 2024; Susongko et al., 2022; Tabacholly et al., 2025). Multiple-choice and essay question formats alone cannot sufficiently promote critical thinking or scientific problem-solving in real contexts (Pujana et al., 2022; Rahmati et al., 2024; Subali et al., 2023).

Furthermore, the analysis of science education assessment tools still relies mainly on Classical Test Theory (CTT), which focuses only on total scores without considering the characteristics of each item or the individual differences among students (Fakhriyah et al., 2022; Toraman & Korkmaz, 2023; Luo & Jia, 2022). As a result, the interpretation of students' literacy performance becomes less diagnostic and fails to represent their actual competencies. The study by Darmawati et al. (2025) also revealed that science literacy instruments for elements, compounds, and mixtures remain conventional, lacking contextual relevance and psychometric precision.

This situation highlights a clear gap in previous studies, as existing instruments have not yet integrated the AKM framework or employed modern measurement models such as the Rasch approach. Consequently, they are limited in accurately assessing students' literacy levels, especially for abstract topics requiring a higher level of conceptual understanding. Therefore, there is a need to develop a contextual, valid, and psychometrically sound science literacy instrument that aligns with AKM characteristics and the updated PISA 2025 indicators. In response to this gap, the present study develops a digital-based AKM science literacy instrument using the Rasch Model to ensure both empirical accuracy and practical applicability in classroom assessment.

Building on this need, advancements in educational technology and psychometric modeling have positioned Rasch Model as alternative to limitations of Classical Test Theory (CTT) (Sunni et al., 2024). This approach measures difficulty level of each item while providing an objective and individual estimate student ability (Chen et al., 2022; Salamah et al., 2022; Susongko et al., 2019). Unlike traditional methods, the Rasch Model provides a diagnostic, scalable, and adaptive framework for assessing science literacy. Developed by George Rasch of Denmark as part of Item Response Theory, this model efficiently calibrates measurement scales, ensuring both psychometric quality of instruments and accuracy of student skill profiles (Olsen, 2003; Wilson, 2023).

Based on this rationale, the present study focuses on developing a Minimum Competency Assessment (AKM) instrument for science literacy on elements, compounds, and mixtures using the Rasch Model for junior high school students. The novelty of this researchin developing a digital-based AKM instrument that integrates Rasch analysis to ensure psychometric precision, contextual relevance, and adaptability. This innovation bridges digital technology and modern psychometrics, producing an instrument that not only measures conceptual knowledge but also evaluates students' ability to apply scientific reasoning and solve real-life problems in accordance with the PISA 2025 framework. Furthermore, this study contributes to competency-based learning and improvement of science education quality by providing valid, adaptive, and technology-integrated tool for assessing students' science literacy in authentic and evidence-based manner.

METHOD

This study focused on developing a science literacy assessment instrument in AKM format on the subject matter of elements, compounds, and mixtures. The research subjects were 175 eighth-grade students at SMP Negeri 1 Juwana who had studied the subject matter of elements, compounds, and mixtures. The researchers used cluster random sampling to select six classes. The study was conducted in May 2024/2025.

The instrument was developed based on the 4D model (Define, Design, Development, Disseminate) presented by Thiagarajan (1974). The Define stage involved a needs analysis through a literature review, interviews with science teachers, and the identification of student characteristics to ensure content's suitability to the learning context. The Design stage included the preparation of question grids based on science literacy indicators, the selection of E-Ujian CBT-based digital media, and the design of instruments in AKM format, which combined objective and subjective questions. Thus, the instruments not only measured knowledge of science concepts, but also the skills to overcome problems in real life.

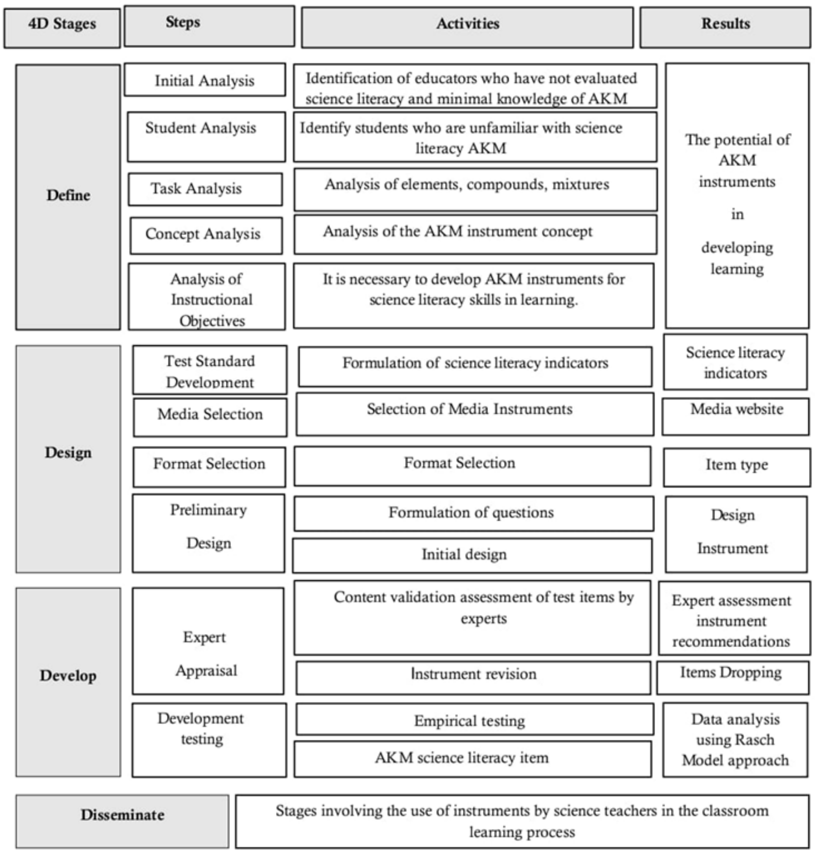


Figure 1. Instrument Development Procedure

The Development stage involved content validation by five science teachers, who served as expert validators. The assessment was analyzed using Aiken's V index (Aiken, 1985) to ensure the relevance, clarity, and appropriateness of the items. The validated instruments were then tested on a limited basis to obtain empirical data. Data analysis was conducted using the Rasch Model approach with the assistance of the Winsteps program. This analysis included item fit tests (infit and outfit MNSQ), item reliability, estimation of question difficulty levels on a logit scale (very easy, easy, difficult, and very difficult), and the preparation of an Item-Person Map to map the distribution of student abilities and the appropriateness of question difficulty levels. The results of this analysis form the basis for determining the suitability of the instrument in measuring science literacy skills comprehensively and objectively.

The Disseminate stage was carried out through the implementation of instruments by science teachers in the classroom learning process. Hence, the developed instruments were not only evaluated through academic testing but also implemented in classroom practices as assessment tools and feedback sources to enhance the quality of learning. The instrument development process based on the 4D model is presented in Figure 1.

RESULT AND DISCUSSION

The results of this study are based on the 4D model development stages (Thiagarajan, 1974), which include Define, Design, Development, and Disseminate. Each stage contributes significantly to the development of the instruments, from needs analysis to classroom implementation. The Define stage is the initial foundation for developing science literacy AKM instruments. The main focus of this stage is to analyze problems and determine the scope and specifications of the product to be developed systematically (Arvianto et al., 2023). The Define stage in this study consists of five components of analysis according to Thiagarajan (1974): initial analysis, student analysis, task analysis, concept analysis, and analysis of instructional objectives.

The initial analysis was conducted through interviews and observations of six science teachers at SMP Negeri 1 Juwana. The results showed that the assessment instruments were used for memorization. Most questions were simple multiple-choice, which could not measure contextual science literacy skills. Analysis of the students revealed that they were unfamiliar with AKM-based question formats, particularly objective-subjective variations that utilize stimuli from real-life concepts. This finding is consistent with Zhang et al. (2023), who emphasized the need for contextual instruments to solve learning needs. The task analysis focused on mapping the material on elements, compounds, and mixtures relevant to the Learning Outcomes of Phase D of the Merdeka Curriculum. The study by Crowe & Bradshaw (2021) confirmed that this material is abstract but has high potential for contextualization. Concept analysis is used to design the structure of questions in the AKM format, considering aspects of science literacy and the context of using personal, socio-cultural, and scientific texts (Guerrero & Torres-Olave, 2022). The Define stage also establishes specific instructional objectives to limit the scope of development and ensure that the item indicators align with the measured science literacy competencies. This formulation refers to the PISA 2025 framework, as suggested by Ulumiyah et al. (2022). Thus, the results of the Define stage confirm the need for an AKM instrument that can comprehensively measure students' science literacy skills.

The design stage serves as a connection between the results of the needs analysis and the instrument construction process in the development stage. According to Streiner et al. (2024), this stage involves four main steps: test standard development, instructional media selection, format selection, and preliminary design. Test standards are developed based on science literacy indicators and learning outcomes for elements, compounds, and mixtures in Phase D of the Merdeka Curriculum. The draft instrument is validated by five science teachers from SMP Negeri 1 Juwana as expert appraisers, using a questionnaire containing item lattices and science literacy skill indicators integrated with the PISA 2025 framework. The instrument was designed in digital format using the E-Ujian CBT platform because it is capable of presenting stimulus-based questions interactively and supports administrative efficiency and data analysis for assessment. This statement is aligned with Istyadji & Sauqina (2023), who emphasized the importance of digital instruments in learning evaluation and the development of educational technology. The AKM format of the science literacy instrument, utilizing the E-Ujian CBT platform, is illustrated in Figure 2.

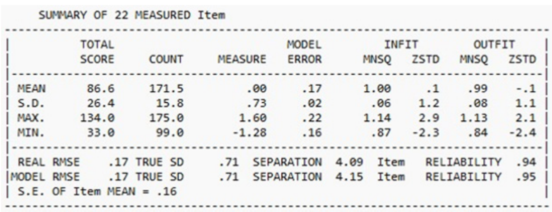
Figure 2. AKM-based science literacy instrument using the E-Ujian CBT platform

The instrument format developed in this study consisted of a combination of objective and subjective tests. The objective tests included multiple-choice, complex multiple-choice, and matching questions, while the subjective tests consisted of short-answer and essay questions. This variety of question formats was designed to measure students' skills in understanding context, organizing ideas, and reasoning scientifically. The result of the design stage was a preliminary instrument consisting of 25 AKM-format science literacy items, which were compiled based on a structured grid and visualized in the form of digital questions. This design became the basis for the development and validity testing in the next stage.

The development stage is a continuous process of instrument design. The primary focus of this stage is to ensure the validity of the instrument's content prior to empirical testing. Validity in this context relates to the extent to which the instrument measures the intended attributes and produces score interpretations appropriate to the measurement objectives, thus requiring adequate scientific evidence (Sumintono & Widhiarso, 2015). Content validity is the first component to be tested, ensuring that each item is relevant to science literacy indicators and learning outcomes. The validation process was performed through expert appraisal by five science teachers from SMP Negeri 1 Juwana. The assessment was conducted using a validation form containing the instrument lattice, science literacy indicators, and assessment aspects covering material suitability, construction accuracy, and language clarity. Content validity analysis was performed using Aiken's V index (Aiken, 1985). Five validators appraised 25 items on a 1–5 Likert scale. Based on Aiken's criteria, the value of V was .87, indicating a high level of suitability; therefore, the instrument was considered suitable for use as a science literacy measurement tool.

According to the results of content validity analysis, 22 of the 25 items were considered valid with an average of .87. Three items (numbers 3, 4, and 5) obtained an Aiken's V score of .80, below the minimum criteria (.87) considered invalid. Further analysis revealed that items 3, 4, and 5 did not fully represent the measured competencies regarding context clarity, stimulus appropriateness, and answer choice accuracy. This item indicated incompatibility with the targeted science literacy indicators. Research by Nurjanah et al. (2023) and Roebianto et al. (2023) emphasizes that items that fail to meet content validity criteria should be eliminated to avoid compromising measurement accuracy and the instrument's effectiveness. Based on these considerations, the three items were found inappropriate for use and were removed from the final instrument. Hence, the instrument contained only 22 items (88%) that were valid and ready for empirical testing. The proportion indicates good content quality and consistency of the instrument in measuring AKM-format science literacy in terms of elements, compounds, and mixtures.

The Development stage is a continuation of the content validation process, which aims to test the instrument directly on the target subjects. The trial was conducted on 175 eighth-grade students at SMP Negeri 1 Juwana, who were selected using a cluster random sampling method. The sample size met the minimum requirements for Rasch Model analysis and the recommendations of Krejcie & Morgan (1970). Science teachers provided additional considerations to ensure the representation of relevant learning contexts. The test data were analyzed using the Rasch Model as part of Item Response Theory. The Rasch Model provides an objective, accurate, and consistent evaluation and refinement of instruments for measuring student ability (Sumintono & Widhiarso, 2015). The analysis focused on four main parameters: item reliability, item fit, item difficulty, and item-person map. The analysis results showed that the item reliability value indicated a high category, which signified the instrument's internal consistency in distinguishing student abilities. This study is aligned with the findings of Avinç & Dogan (2024) who stated that item reliability reflects the stability and consistency of the instrument in assessing abilities. Thus, the developed instrument is deemed to have adequate psychometric characteristics. A summary of the reliability estimation results is shown in Figure 3.



```
SUMMARY OF 22 MEASURED Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT    MEASURE   ERROR       MNSQ  ZSTD   MNSQ  ZSTD   |
|-----------------------------------------------------------------------------|
| MEAN      86.6     171.5        .00     .17       1.00    .1    .99   -.1   |
| S.D.      26.4      15.8        .73     .02        .06   1.2    .08   1.1   |
| MAX.     134.0     175.0       1.60     .22       1.14   2.9   1.13   2.1   |
| MIN.      33.0      99.0      -1.28     .16        .87  -2.3    .84  -2.4   |
|-----------------------------------------------------------------------------|
| REAL RMSE    .17 TRUE SD     .71  SEPARATION  4.09  Item   RELIABILITY  .94 |
|MODEL RMSE    .17 TRUE SD     .71  SEPARATION  4.15  Item   RELIABILITY  .95 |
| S.E. OF Item MEAN = .16                                                     |
-------------------------------------------------------------------------------
```

Figure 3. Reliability Summary based on the Rasch Model

The results of the estimation (see Figure 3) show that the real item reliability value is .94, while the model item reliability value is .95. The reliability value of .94 is included in the "very good" category (.94 < r ≤ 1.00) according to the criteria of Sumintono & Widhiarso (2015). This result analysis confirms that the measurement consistency between items is very high, so the instrument can consistently distinguish students' abilities in science literacy with a proportional distribution of difficulty levels. This result is consistent with Dzin & Lay (2021), who emphasized that item reliability is a crucial parameter to ensure the consistency of instruments in measuring students' competencies.

Furthermore, the model reliability value of .95 indicates that the item structure, as defined by the Rasch Model, has been well established. The item parameter estimates are reliable and are not significantly affected by fluctuations in the sample data. Thus, this instrument is declared to have excellent measurement quality in terms of item consistency and stability, making it suitable for the accurate and reliable assessment of junior high school students' science literacy. Rasch analysis of the AKM-based science literacy instrument reveals that all 22 items that passed content validation have Infit Mean Square and Outfit Mean Square values within the range of .5–1.5 (see Table 1). This range is in accordance with the criteria of Bond & Fox (2015) and Chiu et al. (2025), which confirms that participants' responses to the items are consistent with the model predictions. Thus, from the item fit aspect, the instrument is empirically valid and capable of measuring science literacy competencies stably.

Table 1. Item Analysis Summary Based on the Rasch Model

Item	Item Fit MNSQ (logit)			Item Difficulty (logit)	
	Infit	Outfit	Category	Value	Category
1	.96	.90	Fit	-.64	Easy
2	1.14	1.13	Fit	.31	Difficult
3	.90	.89	Fit	-.56	Easy
4	1.05	1.04	Fit	.24	Difficult
5	.99	1.03	Fit	.60	Difficult
6	1.00	.99	Fit	-.09	Easy
7	.95	.94	Fit	-.35	Easy
8	1.06	1.05	Fit	.54	Difficult
9	.92	.92	Fit	.08	Difficult
10	1.08	1.03	Fit	1.60	Very Difficult
11	.98	.90	Fit	1.60	Very Dif-ficult
12	1.00	.98	Fit	.82	Very Difficult
13	.87	.84	Fit	-.56	Easy
14	1.03	1.03	Fit	-.61	Easy
15	1.02	1.04	Fit	-.64	Easy
16	1.04	1.03	Fit	-.42	Easy
17	1.04	1.03	Fit	-1.28	Very Easy
18	.90	.85	Fit	-.75	Very Easy
19	1.07	1.10	Fit	.16	Difficult
20	1.02	1.03	Fit	-.02	Easy
21	.98	.99	Fit	-.61	Easy
22	1.02	1.00	Fit	.57	Difficult

The analysis results show a varied distribution of item difficulty levels in logit units (Metsämuuronen, 2023). The item difficulty level criteria are presented in Table 2. The standard deviation of .73 indicates that the distribution of questions is proportional, as stated by Sumintono & Widhiarso (2015), so that the instrument can accommodate the abilities of students at various levels. However, several extreme items were found that presented challenges. Two items fell into the very difficult category, with high logit values, so they could only be answered by participants with exceptionally high abilities. This condition risks reducing validity because it is not accessible to most participants, as suggested by Rasyid et al. (2025). On the other hand, there were items with very low difficulty levels that almost all participants answered correctly, so their discrimination capacity was minimal. Following Tennant & Küçükdeveci (2023), items that are too easy do not provide sufficient information to distinguish between participants' abilities.

Table 2. Item Difficulty Level Criteria

Measure the logit item	Category
Measure logit > .73	Very Difficulty
.00 ≤ Measure logit ≤ .73	Difficulty
-.73 ≤ Measure logit ≤ .00	Easy
Measure logit < -.73	Very Easy

(Sumintono & Widhiarso, 2015)

Considering these findings, the instrument is generally of good quality. However, items at the extremes require revision to achieve a more balanced difficulty distribution. This revision is important to ensure that the instrument is not only empirically consistent but also fair, accurate, and capable of providing a comprehensive picture of students' science literacy skills. Thus, items 11 (very difficult) and 17 (very easy) are recommended for revision to ensure the instrument has a more balanced and accurate measurement quality. Details of the items that need revision are shown in Table 3.

Item number 11 has a logit value of 1.60 according to the classification by Sumintono & Widhiarso (2015), which is considered very difficult (logit > .73). A high logit value indicates that only a small number of students were able to answer the question correctly. Based on its content, this item requires an understanding of the environmental impact of ionic compounds, particularly NaCl, which is an advanced concept in chemistry and not necessarily explicitly covered in the junior high school curriculum. Additionally, the true-false format, without adequate contextual stimuli, makes it difficult for students to relate the material to real-life experiences or scientific readings. Technical terms such as "polluting ions" are also relatively unfamiliar to most students, adding to the level of complexity. According to Metsämuuronen (2023), a high level of difficulty is often related to the complexity of the content and the format of the item. Consequently, item 11 requires revision to maintain its assessment of scientific literacy while making it more cognitively and contextually appropriate.

Table 3. Comparison of Original Items and Revised Items

Item (Original)	Item (Revised)
Item Number 11 (Very difficult) Read the following information excerpt! **Disrupted Salt Cycle: A Global Threat**  Development activities and increasing demand for salt are disrupting the stability of the planet's natural salt cycle, affecting the global distribution of salt. The sustainability of ecosystems could be seriously threatened if global salt production and consumption are not regulated in accordance with "The anthropogenic salt cycle." Excessive salt production can disrupt ecosystems. An accelerating salt cycle can release these ions in higher doses than usual. The release of these ions can cause environmental problems. Source: nationalgeographic. grid.id	**Item Number 11 (Very difficult)** Read the following information excerpt! **Disrupted Salt Cycle: A Global Threat**  Development activities and increasing demand for salt are disrupting the stability of the planet's natural salt cycle, affecting the global distribution of salt. The sustainability of ecosystems could be seriously threatened if global salt production and consumption are not regulated in accordance with "The anthropogenic salt cycle." Excessive salt production can disrupt ecosystems. An accelerating salt cycle can release these ions in higher doses than usual. The release of these ions can cause environmental problems. Source: nationalgeographic. grid.id

Choose True (T) or False (F) from the statements

Statement	T	F
Salt consists of cations and anions		
Salt can contain calcium, magnesium, potassium, and sulfate ions.		
Sodium and chloride ions are found in seawater.		
Salt consists of sodium and chloride ions.		

Item Number 17 (Very Easy)
Read the passage and examine the following table to answer the question!

The Effect of Alum on Clarifying Turbid Water

Andi is a researcher who wants to know the effectiveness of alum in clarifying turbid water. To pursue his interest, he experimented by preparing three containers filled with the same volume of turbid water. Andi experimented for 24 hours. After 24 hours, he observed and recorded the level of turbidity in each container, as shown in the following table.

Observation Data of the Effect of Alum on the Clearing of Cloudy Water

No	Mass (gr)	Result
1	1	Clear
2	2	Medium cloudy
3	0	Very cloudy

Explain the correct conclusion to answer Andi's initial experiment objective based on the data obtained from the experiment and the reason why alum can purify water. Write your answer in the column below!

Choose True (T) or False (F) from the statement

Statement	T	F
Salt consists of cations and anions		
Salt can contain calcium, magnesium, potassium, and sulfate ions.		
Sodium and chloride ions are found in seawater.		
Salt consists of sodium and chloride ions.		

Item Number 17 (Very Easy)
Read the passage and examine the following table to answer the question!

The Effect of Alum on Clarifying Turbid Water

Andi conducted an experiment to determine the effectiveness of alum in purifying turbid water. He prepared three containers filled with turbid water of equal volume. After 24 hours, he recorded the turbidity levels of the water as follows.

Observation Data of the Effect of Alum on the Clearing of Cloudy Water

No	Mass (gr)	Result
1	1	Clear
2	2	Medium cloudy
3	0	Very cloudy

Based on the data from Andi's experiment, explain the conclusions. Compare the results in the three containers and provide a scientific explanation of why alum can affect the level of turbidity in water. Include the reason why higher levels of alum do not always produce clearer water. Write your answer in the column below!

Item number 17 has a logit value of -1.28, which falls into the very easy category (logit < .73). Almost all students were able to answer it correctly, so this item is ineffective in distinguishing ability levels. Item formats that only require simple conclusions without the need for scientific analysis or reasoning tend to be less suitable for AKM-format science literacy instruments. According to Noviantini et al. (2023), items that are too simple need to be revised to require scientific reasoning skills, so that they remain relevant to the objectives of science literacy assessment. Therefore, revision of item number 17 is necessary to enable the instrument to measure students' abilities more comprehensively.

Classification of item difficulty levels based on logit values from the Rasch analysis to assess items that can differentiate student abilities. The interpretation results are reinforced with the visualization of the item-person map as found by Aryadoust & Sayama (2021) to show the relative distribution between student skills and item difficulty levels. Based on this map, the majority of students are within the logit range of 0 to

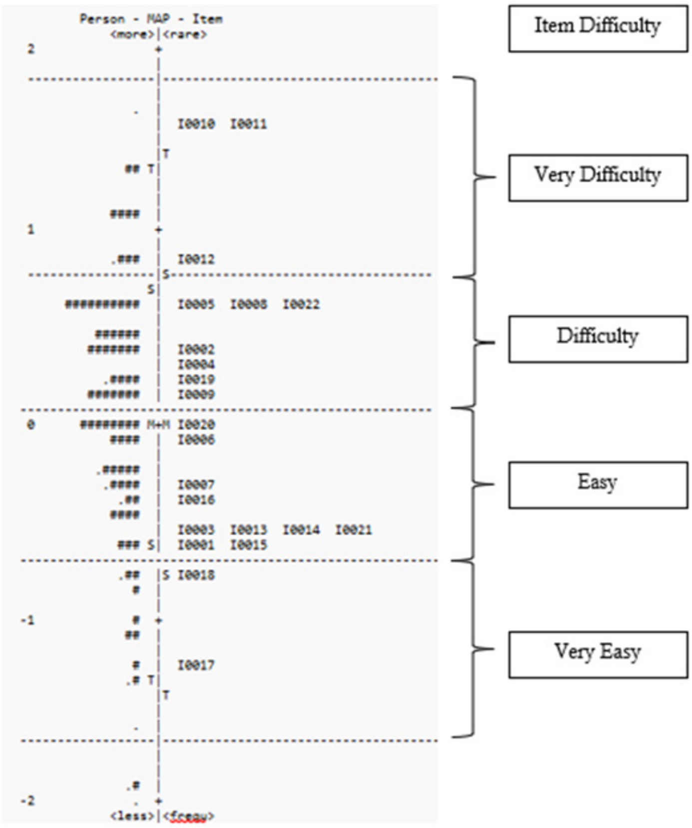


Figure 4. Item-Person Map

+1, indicating that the ability of students at SMP Negeri 1 Juwana to answer the item category is easy to difficult. Meanwhile, the items are equally distributed from logit –2 to +2, which represents a range of difficulty from very easy to very difficult, as shown in Figure 4. The results of this study confirm that the instrument strikes a good balance between the level of difficulty of the items and the students' ability to demonstrate a variety of science literacy skills optimally. The map shown in Figure 4 provides an overview of the compatibility of the instruments with the science literacy profiles of junior high school students.

Finally, the dissemination stage demonstrated that the 22 valid items were effectively implemented by science teachers at SMP Negeri 1 Juwana as learning instruments for the topic of elements, compounds, and mixtures. The integration of this AKM-based science literacy instrument into classroom practice strengthened students' ability to connect scientific concepts with real-life contexts and encouraged teachers to apply more contextual, competency-oriented assessments. These findings highlight the potential of the Rasch Model to enhance both measurement accuracy and pedagogical relevance in science education.

This study [A5.1]contributes to educational research by providing empirical evidence on the effectiveness of the Rasch Model in developing valid and adaptive assessment instruments aligned with national (AKM) and international (PISA 2025) standards. The findings suggest that psychometric-based assessments can bridge the gap between curriculum policy and classroom practice by supporting evidence-based evaluation of students' science literacy competencies. Practically, this instrument serves as a model for teachers to design contextual assessment tools that not only measure knowledge but also foster critical thinking and problem-solving skills in authentic learning situations.

However, this study is limited to the topic of elements, compounds, and mixtures and was implemented within a single school context. Therefore, future research is recommended to extend the development of AKM-based science literacy instruments to other scientific domains and broader educational settings. Such expansion will strengthen the generalizability and applicability of the Rasch-based assessment model, providing deeper insights into how competency-oriented assessments can enhance the quality of science learning and policy implementation in Indonesia.

CONCLUSION

The study concluded that the AKM-based science literacy instrument, developed using the Rasch Model, demonstrated strong validity and reliability in assessing elements, compounds, and mixtures. The 22-item instrument effectively measures junior high school students' science literacy skills with high accuracy and consistency. The Rasch-based approach offers a psychometrically sound framework for developing contextual and competency-oriented assessments that align with AKM and PISA 2025 standards. Overall, the instrument offers a meaningful contribution to strengthening science education assessment practices and promoting students' critical and scientific reasoning skills through authentic evaluation.

REFERENCES

Agnesa, O. S., Hanidar, E., Afifi, N., & Rahmadana, A. (2023). Pelatihan Pembelajaran Literasi Sains Menggunakan Keterampilan Berpikir Tingkat Tinggi Sebagai Penyesuaian terhadap Instrumen Assesmen Kompetensi Madrasah Pengetahuan, Budaya dan Pengalaman Sebelumnya untuk Mengembangkan Assesmen Komptensi Literasi Diisi. *I-Com: Indonesian Community Journal*, 3(1), 88–102.

Aiken, L. R. (1985). Three Coefficients for Analyzing The Reliability and Validity of Ratings, Educational and Psychological Measurement. *Educational and Psychological Measurement*, 45(1), 131–142.

Alwalya, G., & Rendy, D. B. (2024). Pengembangan E-Magazine Tema Lontong Kupang Sidoarjo pada Materi Unsur, Senyawa dan Campuran. *UNESA Journal of Chemical Education*, 13(3), 226–237.

Annisa, D., Sutrisno, H., Laksono, E. W., & Yanda, S. N. (2024). Evaluating Students' Academic Resilience in Chemistry Learning: Insights from a Rasch Model Analysis. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 6(3), 328–349.

Arvianto, Slamet, S. Y., & Andayani, A. (2023). Designing an Instrument to Measure Digital Literacy Competence Using the 4D Model. *International Journal of Instruction*, 16(4), 845–860.

Aryadoust V., Ng, L. Y., & Sayama, H. (2021). A Comprehensive Review of Rasch Measurement in Language Assessment: Recommendations and Guidelines for Research. *Language Testing*, 38(1), 6–40.

Avinç, E., & Dogan, F. (2024). Digital Literacy Scale : Validity and Reliability Study with the Rasch Model. *Education and information Technologies*, 29(17), 22895-22941

Ayuningtyas, V. E. H., Viyanti, & Setyarini, M. (2023). Teachers ' Perception of the Minimum Competency Assessment Instrument (AKM) based on Lynk.id to Measure Students ' Scientific Literacy on Solar System Material. *Jurnal Penelitian Pendidikan IPA*, 9(11), 9883–9892.

Bond, T. G., & Fox, C. (2015). *Applying the rasch model: fundamental measurement in the humansciences (3rd ed.)* New Jersey: Lawrence Erlbaum Associates Lawrence Erlbaum Associates.

Broderick, N. (2025). Exploring Different Visions of Scientific Literacy in Irish Primary Science Education: Core Issues and Future Directions. *Irish Educational Studies,* 44(1), 73–93.

Chen, Y.-F., Cannady, M. A., Fauville, G., Strang, C., & Berkeley, U. (2022). Working Toward an International Assessment of Ocean Literacy: Validating Instrument With Rasch Measurement Model. *Educational Research Association (AERA) Annual Meeting*, 1(1), 56–57.

Chiu, E. C., Lee, Y., Wu, Y., & Lee, S. (2025). Item Y.-F.Analysis of the Behavior Rating Inventory of Executive Function Adults Version Using the Rasch Model in Two Forms for People with Schizophrenia in Taiwan. *International Journal of Methods in Psychiatric Research*, 34(3), 1–7.

Crowe, J., & Bradshaw, T. (2021). *Chemistry for the biosciences: the essential concepts*. England: Oxford University Press.

Darmawati, V. D., Wulandari, A. Y. R., Putra, D. B. R. A., Endang, & Sulastri1. (2025). Validitas Instrumen Soal Literasi Sains pada Konteks Batik Madura Materi Unsur, Senyawa, dan Campuran. *Jurnal Pendidikan*, 14(September), 723–731.

Deta, U. A., A, S. K., Laila, L., Prahani, B. K., & Suprapto, N. (2025). PISA Science Framework 2018 vs 2025 and Its Impact in Physics Education : Literature Review. *Momentum: Physics Education Journal*, 8(1), 95–107.

Ducha, N., & Husna, U. H. (2025). Development of Assessment Instruments to Measure Science Literacy Skills 11 th of Grade High School Students on Excretory System. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, 14(1), 98–109.

Dwiwulandari, A., Hakim, M., & Hasnawati. (2024). Analisis Kemampuan Literasi Sains Siswa Kelas V pada Pembelajaran IPAS SDN 1 Jagaraga Indah. *Jurnal Pendidikan, Sains, Geologi, dan Geofisika*, 5(4), 799-807.

Eralita, N., & Setiawan, Y. A. (2022). Studi Literasi Kimia Siswa Sekolah Menengah Kejuruan pada Materi Zat dan Perubahan. *Jurnal Eduscience,* 9(3), 792–801.

Faisal, S., Rusilowati, A., & Susilaningsih, E. (2023). Science Literacy Assessment Instrument for Additives and Addictive Substances: Development, Validation, and Rasch Model Analysis. *Jurnal Penelitian Pendidikan IPA*, 9(10), 7826–7836.

Fakhriyah, F., Rusilowati, A., Nugroho, S. E., Saptono, S., Ridlo, S., Mindyarto, B., & Susilaningsih, E. (2022). The Scientific Argumentative Skill Analysis Reviewed from the Science Literacy Aspect of Pre-Service Teacher. *International Journal of Evaluation and Research in Education*, 11(4), 2129–2139.

Guerrero, G. R., & Torres-Olave, B. (2022). Scientific Literacy and Agency Within the Chilean Science Curriculum: A Critical Discourse Analysis. *Curriculum Journal,* 33(3), 410–426.

Hawariah, Mutmainah, P. A., & Agustina, S. (2024). Analisis Ketercapaian Kemampuan Asesmen Kompetensi Minimum (AKM) melalui Literasi Sains Siswa SMPN 5 Kota Bima. *Jurnal Redoks: Jurnal Pendidikan Kimia dan Ilmu Kimia,* 7(2), 80–87.

Hudson, A. K., Moore, K. A., Han, B., Wee Koh, P., Binks-Cantrell, E., & Malatesha Joshi, R. (2021). Elementary Teachers' Knowledge of Foundational Literacy Skills: A Critical Piece of the Puzzle in the Science of Reading. *Reading Research Quarterly,* 56(S1), S287–S315.

Istyadji, M., & Sauqina. (2023). Conception of Scientific Literacy in the Development of Scientific Literacy Assessment Tools: A Systematic Theoretical Review. *Journal of Turkish Science Education*, 20(2), 281–308.

Kim, S. L., & Kim, D. (2021). English Learners' Science-Literacy Practice through Explicit Writing Instruction in Invention-Based Learning. *International Journal of Educational Research Open*, 2, 100029.

Krejcie, R. V., & Morgan, D. W. (1970). Sample Size Determination Table. *Educational and Psychological Measurement*, 30(3), 607-610.

Lubis, M. A. R., Alvina, S., Muliaman, A., Lukman, I. R., & Fakhrah, F. (2024). The Effect of the Read, Answer, Discuss, Explain, And Create (Radec) Learning Model on Students' Chemical Literacy Skills in the Independent Curriculum. *Chimica Didactica Acta (Journal of Chemistry & Chemistry Education)*, 12(02), 118–123.

Luo, X., & Jia, H. (2022). When Scientific Literacy Meets Nationalism: Exploring the Underlying Factors in the Chinese Public's Belief In COVID-19 Conspiracy Theories. *Chinese Journal of Communication*, 15(2), 227–249.

Mazidah, L. N., & Suwarna, I. P. (2025). Meta Analisis Kemampuan Literasi Sains pada Mata Pelajaran Fisika. 10(1), 381–388.

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

Mohd Dzin, N. H., & 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.

Noviantini, N. M. H., Sariyasa, & Ardana, M. (2023). Development of Minimum Competency Assessment (AKM) Instruments to Measure The Numeration Ability of Seventh Grade of Junior High School Students. *International Journal of Trends in Mathematics Education Research,* 6(1), 65–71.

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), 2023–2125.

OECD. (2023). *PISA framework kerangka kerja sains 2025*. Publikasi OECD.

OECD. (2025). *PISA 2025 framework for science, reading, and mathematics.* OECD Publishing.

Olsen, L. W. (2003). *Essays on georg rasch and his contributions to statistics*. Denmark: Københavns Universitet.

Permatasari, A. D. A., Handayani, R. D., Siswati, B. H., & Bachtiar, A. A. (2023). Effectiveness Test Digital Assessment based Moodle as a Digital Assessment Instrument for Measuring Science Literacy Skills for Junior High School Students. *JPPS (Jurnal Penelitian Pendidikan Sains),* 13(1), 220–232.

Pratama, R., Alamsyah, M., Siburian, M. F., Marhento, G., Jonathan, G. L., & Susanti, W. (2024). Analisis Kemampuan Literasi Sains Peserta Didik pada Mata Pelajaran IPA. *Jurnal Pendidikan MIPA*, 14, 576–581.

Pujana, L. A., Dwijayanti, I., & Siswanto, J. (2022). Pengembangan Bahan Ajar Berbasis Model Pembelajaran Clis Seri AKM untuk Meningkatkan Kemampuan Literasi Sains Siswa SD. *Pendas : Jurnal Ilmiah Pendidikan Dasar*, 7(2), 589–604.

Rahmati, U., Ginting, F. W., Novita, N., Syafrizal, & Sakdiah, H. (2024). Pengembangan Instrumen Tes Fisika Berbasis Literasi Sains pada Materi Gerak Melingkar. *GRAVITASI: Jurnal Pendidikan Fisika dan Sains*, 7(01), 10–21.

Rasyid, F., Istiyono, E., Gunawan, C. W., & Kijambu, J. B. (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.

Roebianto, A., Savitri, S. I., Aulia, I., Suciyana, A., & Mubarokah, L. (2023). Content Validity: Definition and Procedure of Content Validation in Psychological Research. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 30(1), 5–18.

Salamah, A., Prastowo, T., & Hariyono, E. (2022). The Development Research : Assessment Instruments of Science Literacy Based on Minimum Competency Assessment (AKM) Level 4 th for Solar System Concepts. *Jurnal Penelitian Dan Pengembangan Pendidikan Fisika*, 8(1), 115–124.

Streiner, D. L., Norman, G. R., & Cairney, J. (2024). *Health measurement scales--a practical guide to their development and use*. England: Oxford University Press.

Subali, B., Faizah, Z., & Sidiq, M. (2023). Indonesian National Assessment Support : Can RE-STEM

Android App Improve Students ' Scientific Literacy Skills. *International Journal of Evaluation and Research in Education (IJERE),* 12(3), 1399–1407.

Sumintono, B., & Widhiarso, W. (2015). *Pemodelan rasch pada assessment pendidikan*. Cimahi: Trim Komunikata Publishing House.

Sunni, F. M., Dawana, I. R., Jannah, R., Sunarti, T., Madlazim, Wasis, & Rizki, I. A. (2024). Research Trend of Rasch Model in the Development of Scientific Literacy Assessment Through Bibliometrics. *JIPI (Jurnal IPA dan Pembelajaran IPA),* 8(4), 413–429.

Susongko, P., Widiatmo, H., Kusuma, M., & Afiani, Y. (2019). Development of Integrated Science-Based Science Literacy Skills Instruments using the Rasch Model. *Unnes Science Education Journal,* 8(3), 277–292.

Susongko, P., Yuenyong, C., & Zainudin, A. (2022). Buddhist Critical Thinking Assessment using Rasch Model. *Kasetsart Journal of Social Sciences*, 43(2), 285–292.

Tabacholly, F. A., Susongko, P., & Nafiati, D. A. (2025). Construct Validation of Science Literacy Instrument with Rasch Modeling on Students of Grade VIII Junior High School. *Journal of English Language and Education*, 10(1), 378–388.

Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch Measurement Model in Rehabilitation Research and Practice: Early Developments, Current Practice, and Future Challenges. *Frontiers in Rehabilitation Sciences*, 4, 1208670.

Thiagarajan. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Minnesota: University of Minnesota.

Toraman, Ç., & Korkmaz, G. (2023). What is the "Meaning of School" to High School Students? A Scale Development and Implementation Study Based on IRT and CTT. *SAGE Open*, 13(3), 1–15.

Ulumiyah, D., Sumantri, M. S., Rahmawati, Y., & Iasha, V. (2022). An Analysis of Science Literacy Ability Elementary School Students. *Journal Basicedu*, 6(3), 3544–3553.

Valentini, N. C., Duarte, M. G., Zanella, L. W., & Nobre, G. C. (2022). Test of Gross Motor Development-3: Item Difficulty and Item Differential Functioning by Gender and Age with Rasch Analysis. *International Journal of Environmental Research and Public Health*, 19(14), 8667-8684.

Valladares, L. (2021). Scientific Literacy and Social Transformation: Critical Perspectives about Science Participation and Emancipation. *Science & Education*, 30(3), 557–587.

Waini, T. S. J., Primairyani, A., Wardana, R. W., Uliyandari, M., & Sutarno, M. (2024). Analisis Kemampuan Literasi Sains Siswa Kelas VIII di SMPN 9 Kota Bengkulu pada Isu Pencemaran Lingkungan di Kota Bengkulu. *DIKSAINS: Jurnal Ilmiah Pendidikan Sains*, 5(1), 62–69.

Wilson, M. (2023). Constructing measures: an item response modeling approach. *Canada: Routledge*.

Zhang, L., Liu, X., & Feng, H. (2023). Development and Validation of an Instrument for Assessing Scientific Literacy from Junior to Senior High School. *Disciplinary and Interdisciplinary Science Education Research*, 5(1), 21-36.