HYDROGEN
JURNAL KEPENDIDIKAN KIMIA
https://e-journal.undikma.ac.id/index.php/hydrogen/index

February 2026. 14(1)
e-ISSN: 2656-3061
p-ISSN: 2338-6487
pp. 33-42

Validation of Ethnochemistry-Integrated Two-Tier Items for Chemical Literacy among Pre-service Chemistry Teachers

Dyah Puspitasari Ningthias*, Ermia Hidayanti, Yunita Arian Sani Anwar, Nora Listantia, Dodiy Firmansyah, Baiq Viona Olganita Sugondo
Chemistry Education, Faculty of Teacher Training and Education, Mataram University, Jl. Majapahit No. 62, Mataram, NTB, Indonesia
*Corresponding Author email: dyahpuspitasari@staff.unram.ac.id

Article History
Received: 04-02-2024
Revised: 09-02-2026
Published: 26-02-2026

Keywords: Chemical Literacy; Ethnochemistry; Instrumen Validation; Pre-service chemistry teacher; Two-tier items.

Abstract
Chemical literacy is an essential competency for preservice chemistry teachers because it supports the ability to understand chemical concepts, interpret chemical information, and apply chemistry in daily and socio-scientific contexts. However, chemical literacy assessment is often limited to conventional multiple-choice tests that do not adequately measure reasoning or contextual application. This study aimed to validate ethnochemistry-integrated two-tier multiple-choice with 10 items test for measuring chemical literacy among preservice chemistry teachers. A quantitative survey design was employed with 35 participants from a public university in West Nusa Tenggara selected through convenience sampling. This approach allowed the researchers to obtain initial empirical evidence on the psychometric quality of the ethnochemistry-integrated two-tier instrument for measuring chemical literacy among pre-service chemistry teachers. The instrument was developed by adapting chemical literacy indicators and integrating local ethnochemical contexts into two-tier items, followed by expert judgment validation. Data were collected through face-to-face, paper-based administration and analyzed using the Rasch measurement model to examine item difficulty calibration, item fit statistics, reliability, separation indices, and unidim The Wright map shows that most pre-service chemistry teachers were clustered around the mean ability range (−1 to +1 logits), indicating moderate chemical literacy, while the items were well distributed in difficulty with more challenging tasks such as Item 8 and Item 4 located at higher logits. Rasch analysis further confirmed acceptable measurement quality, with person reliability of 0.71 and Cronbach's alpha of 0.74, item fit statistics mostly within the acceptable range (Infit MNSQ 0.78–1.43), and strong unidimensionality evidenced by 40.9% variance explained by the measures.ensionality through residual analysis. This study is novel in validating an ethnochemistry-integrated two-tier multiple-choice instrument using Rasch analysis to provide a culturally responsive and reasoning-based assessment of chemical literacy among pre-service chemistry teachers by connecting chemical concepts with local wisdom and real-life contexts.

How to Cite: Ningthias, D. P., Hidayanti, E., Anwar, Y. A. S., Listantia, N., Firmansyah, D., & Sugondo, B. V. O. (2026). Validation of Ethnochemistry-Integrated Two-Tier Items for Chemical Literacy among Pre-service Chemistry Teachers. *Hydrogen: Jurnal Kependidikan Kimia*, *14*(1), 33–42. https://doi.org/10.33394/hjkk.v14i1.19581

https://doi.org/10.33394/hjkk.v14i1.19581

INTRODUCTION

Chemical literacy is a crucial competency in science education because it equips individuals with the ability to understand chemical concepts, interpret chemical information, and apply chemical knowledge to real-life situations. In modern society, chemistry-related issues are embedded in daily decision-making, such as evaluating food additives, understanding medicine labels, selecting household cleaning products, and responding to environmental concerns like pollution and waste management (Pebriani & Aini, 2025; Pratama, et. al., 2025;

Sjöström, et al., 2024). Chemical literacy goes beyond memorizing formulas or reactions; it involves reasoning scientifically, using evidence to explain phenomena, and communicating chemical ideas in meaningful ways. It also supports informed citizenship, enabling people to participate in public discussions about science-based policies, health, and sustainable development. Furthermore, chemical literacy encourages critical thinking and problem-solving skills needed to face global challenges, including climate change, energy production, and water treatment. Therefore, developing chemical literacy is essential for building scientifically responsible individuals who can make rational choices and contribute positively to society (Alwatoni, et al., 2024; Dewi, et al., 2024; Hany & Purwaningsih, 2025).

For preservice chemistry teachers, chemical literacy is particularly important because they are future educators responsible for shaping students' understanding of chemistry and its relevance to everyday life. Teachers with strong chemical literacy are better prepared to explain chemical concepts clearly, connect abstract ideas to real-world contexts, and correct misconceptions effectively. Chemical literacy also helps teachers design learning experiences that promote inquiry, reasoning, and scientific communication rather than rote learning (Wardani, et al., 2024; Yuendita & Rohaeti, 2025).

In addition, preservice teachers must be able to interpret scientific information critically, especially when dealing with socio-scientific issues such as environmental pollution, food safety, and health-related chemical products. Without adequate chemical literacy, teachers may struggle to guide students in applying chemistry to practical problems. Thus, strengthening chemical literacy among preservice chemistry teachers is essential for improving chemistry instruction quality and fostering scientifically literate future generations (Manik, 2024; Marfuatun, et al., 2024; Nugraheni & Srisawasdi, 2025).

Ethnochemistry-integrated two-tier multiple-choice questions are valuable because they assess chemical literacy while incorporating culturally relevant contexts that students recognize from everyday life. Ethnochemistry connects chemical concepts to local traditions, materials, and practices such as fermentation, herbal medicine, traditional food processing, or cultural rituals. This contextualization enhances meaning-making and helps learners relate chemistry to their lived experiences (Ashari & Munawwarah, 2025; Natania, et al., 2025).

Meanwhile, the two-tier multiple-choice format strengthens assessment quality by measuring not only students' answers (first tier) but also their reasoning (second tier). This structure allows researchers to identify misconceptions, evaluate conceptual understanding, and distinguish guessing from genuine comprehension. Compared to conventional multiple-choice tests, ethnochemistry-integrated two-tier questions provide richer diagnostic information and better reflect the multidimensional nature of chemical literacy, which includes understanding concepts, applying them in context, and explaining scientific reasoning (Junaidi, et al., 2025; Ningtyas, et al., 2025; Pamungkas, et al., 2024).

A valid assessment instrument is essential to accurately measure chemical literacy using ethnochemistry-integrated two-tier multiple-choice questions. This instrument combines cultural contexts and scientific reasoning, poor validity can lead to misleading interpretations of students' abilities. Validity ensures that each item truly represents chemical literacy indicators, such as conceptual understanding, application of chemistry in daily life, and reasoning based on evidence Hasanah, et al., 2025).

It also ensures that ethnochemical contexts do not introduce bias or irrelevant difficulty that could disadvantage certain respondents. In two-tier items, validity is especially important because both the answer and justification must function appropriately; if the reasoning options are unclear or not aligned with the intended concept, results may reflect confusion rather than literacy (Ningroom, et al., 2025). A valid instrument supports accurate diagnosis of misconceptions and provides trustworthy data for curriculum development and teacher education. Therefore, establishing validity is necessary to ensure the instrument produces meaningful, fair, and interpretable measurement outcomes.

The Rasch model is widely used in educational measurement because it provides strong psychometric evidence for validating assessment instruments, including ethnochemistry-integrated two-tier multiple-choice questions.

Unlike classical test theory, Rasch analysis evaluates item quality and respondent ability on the same logit scale, enabling detailed examination of item difficulty and person competence (Adam, et al., 2025; Hidayat, et al., 2025).

This model allows researchers to test whether items fit the expected measurement pattern, ensuring that responses reflect the targeted construct of chemical literacy. Rasch analysis also provides reliability and separation indices, which indicate how well the instrument distinguishes different ability levels and how stable item difficulty estimates are across samples. Additionally, the Rasch model supports unidimensionality testing through residual analysis, ensuring that the instrument measures a single dominant trait. By identifying misfitting items and examining response consistency, Rasch modeling strengthens the validity, fairness, and interpretability of the instrument for research and educational practice (Hidayat, et al., 2021; Jumriani, et al., 2025).

Despite the growing emphasis on chemical literacy as a key outcome of science education, assessment practices in chemistry still rely heavily on conventional multiple-choice tests that mainly capture factual recall rather than students' reasoning and contextual application. Although two-tier multiple-choice instruments have been developed to diagnose misconceptions and evaluate conceptual understanding, most existing tools are designed using decontextualized chemical situations and rarely integrate culturally meaningful contexts. In addition, ethnochemistry-based learning has increasingly been promoted to connect chemistry with local wisdom and everyday practices; however, empirical studies focusing on *assessment* tools grounded in ethnochemistry remain limited. More importantly, few validated instruments specifically target pre-service chemistry teachers, even though they are critical agents for fostering chemical literacy in future classrooms. Many previous validations also depend on classical test theory, which provides limited information on item functioning across ability levels.

As a result, current chemical literacy assessments do not adequately capture how pre-service teacher reason with chemical ideas in culturally meaningful contexts, nor do they provide robust measurement evidence about item performance across different ability levels. Therefore, there is a need to develop and validate ethnochemistry-integrated two-tier items using Rasch measurement to ensure accurate and fair chemical literacy assessment.

METHOD

Research Design

This study adopted a quantitative survey design because the main purpose was to validate an ethnochemistry-integrated two-tier instrument for measuring chemical literacy among pre-service chemistry teachers using numerical evidence. A survey method was considered appropriate since validation requires collecting responses from a relatively large group in a standardized manner to evaluate item performance objectively. The two-tier format was selected to capture not only participants' answer choices but also the reasoning behind those choices, allowing the identification of misconceptions and levels of scientific justification. Integrating ethnochemistry into the items required empirical testing to ensure that culturally contextualized content remained measurable, unbiased, and interpretable across respondents (Fischer, et al., 2023; Jamieson, et al., 2023).

Quantitative analysis of survey data using the Rasch measurement model enabled systematic examination of item functioning, including item difficulty calibration, person ability estimation, and item fit statistics, providing defensible evidence of construct validity and reliability. The Rasch model also supported evaluation of unidimensionality and separation indices, ensuring that the ethnochemistry-integrated two-tier instrument operates consistently and measures chemical literacy accurately among pre-service chemistry teachers (Qudratuddarsi, et al., 2024; Schoonenboom, 2023).

Research Subject

Convenience sampling was used because the research focused on initial instrument validation with accessible respondents who matched the target characteristics and were available during the data collection period. This technique allowed efficient recruitment within limited time and resources while ensuring participants had sufficient background in chemistry education to respond meaningfully to the two-tier items (Golzar, et al., 2022; Qudratuddarsi, et al., 2025).

The participants consisted of 35 pre-service chemistry teachers enrolled in a public university in West Nusa Tenggara, selected due to their relevance as prospective chemistry educators and their exposure to chemistry concepts and local cultural contexts reflected in the instrument.

Instrument Development

The instrument development process began by adjusting chemical literacy indicators to align with the research context and the integration of ethnochemistry, ensuring the construct reflected both core chemistry understanding and culturally relevant applications. Indicators were adapted from established chemical literacy frameworks and refined to emphasize conceptual knowledge, interpretation of chemical phenomena, and reasoning in everyday and local cultural practices.

Based on the indicators, ethnochemistry-integrated two-tier items were constructed, with the first tier assessing content-related answers and the second tier probing the justification for each choice to capture reasoning quality. Item writing incorporated local ethnochemical examples from West Nusa Tenggara to enhance contextual relevance while maintaining scientific accuracy. The initial pool of items was then validated through expert judgment involving chemistry education experts and practitioners familiar with ethnochemistry. Experts reviewed each item for content relevance, clarity, alignment with indicators, cultural appropriateness, and scientific correctness, and their feedback guided item revision prior to pilot administration (Murashko, et al., 2025).

Data Collection and Analysis

Data were collected through face-to-face administration to ensure uniform testing conditions and to reduce incomplete responses. The ethnochemistry-integrated two-tier items were administered in a supervised classroom setting using a paper-based examination format. This approach was selected to provide equal access for all participants, avoid technical constraints related to internet connectivity or device availability, and allow participants to focus on the contextual ethnochemistry scenarios embedded in each item.

Before the test began, participants were given standardized instructions on how to respond to both tiers, including selecting an answer in the first tier and choosing the most appropriate reasoning option in the second tier (Kassiavera, et al., 2024).

Data analysis employed the Rasch measurement model to examine the psychometric quality of the ethnochemistry-integrated two-tier instrument. Responses were coded and entered into Rasch analysis software to estimate item difficulty and person ability on a common logit scale (Peeters & Augustine, 2023). The analysis included examination of item-person distribution through a Wright map, reliability and separation indices for persons and items, and internal consistency using Cronbach's alpha (Post, et al., 2022). Item fit was evaluated using Infit and Outfit mean square (MNSQ) statistics, standardized fit indices (ZSTD), and point-measure correlations to confirm alignment with the intended cons truct. Unidimensionality was assessed using principal component analysis (PCA) of standardized residuals Alnahdi, et.al., 2025).

RESULTS AND DISCUSSION

Wright Map

The Wright map is a visual representation produced through Rasch model analysis that places respondents' abilities and item difficulties on the same logit scale, allowing direct comparison between learner competence and assessment demands (Ismail, et al., 2021; Sutrisno, et al., 2025). In this study, the left side of the map illustrates the distribution of preservice chemistry teachers' chemical literacy abilities, while the right side shows the difficulty levels of the ethnochemistry-integrated two-tier multiple-choice items. Most respondents are clustered around the mean ability level, approximately between −1 and +1 logits, indicating moderate chemical literacy among most participants.

A smaller number of respondents appear at higher and lower logit positions, reflecting variation in students' capacity to interpret chemical concepts embedded in cultural contexts such as traditional practices, fermentation, equilibrium, and colligative properties. From the item distribution, the Wright map demonstrates a well-spread range of difficulty levels across the instrument, which is consistent with Rasch studies showing that a balanced spread of item difficulty across the logit scale indicates appropiate person-item targeting (Tennant & Conaghan, 2023).

Items positioned at higher logits, such as item 8 and item 4, indicate more demanding tasks that require higher-order reasoning, including pH calculation, equilibrium shifts, and quantitative analysis within ethnochemical contexts like *nginang/nyirih*. In this perspective, the discrimination of high-logit items does not merely indicate that they are more difficult, but also reflects how the Wright map aligns item difficulty with respondent ability, so that only participants with higher Rasch measures tend to solve these items succesfully, while lower-ability participants are more likely to struggle with them (Post, et al., 2022). Items located near or below the mean, including item 1 and item 6, represent more accessible questions focusing on foundational concepts such as acid–base behavior, ion interactions, and basic stoichiometry linked to familiar cultural practices. This distribution illustrates how the items function across varying levels of chemical literacy without restating evaluative conclusions.

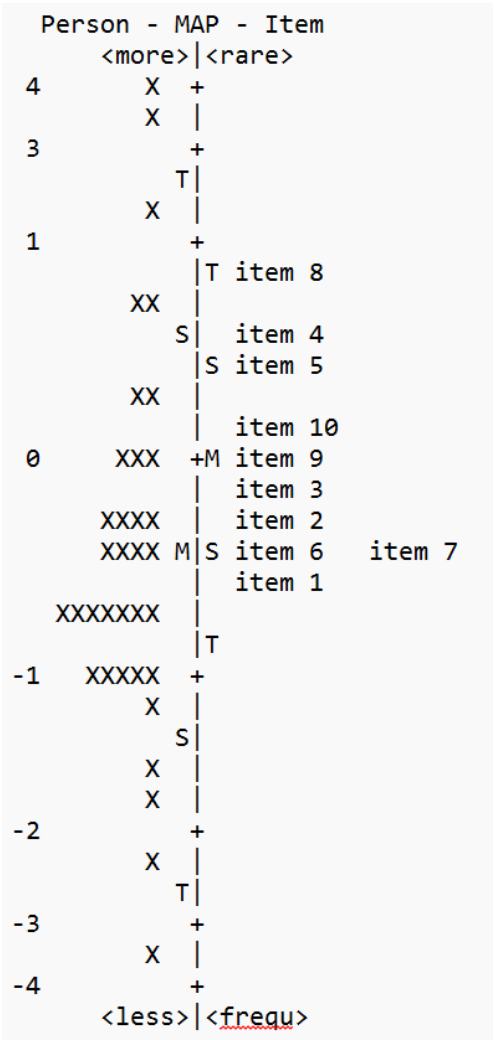

Figure 1. Wright Map

Reliability and Separation

In Rasch model analysis, reliability and separation indices are used to evaluate the consistency of measurement and the instrument's ability to distinguish between different levels of respondent ability and item difficulty. Person reliability reflects how consistently the instrument differentiates respondents based on chemical literacy, while item reliability indicates the stability of item difficulty estimates across samples (Medvedev & Krägeloh, 2025).

Separation indices describe how well respondents or items can be classified into statistically distinct strata, which can be estimated using the formula H (separation) = {(4 x separation) + 1}/3. In this study, the person reliability of 0.71 and Cronbach's alpha of 0.74 indicate acceptable internal consistency in measuring preservice chemistry teachers' chemical literacy. The person separation of 1.98 corresponds to approximately three ability strata ([H = (4 x 1.98 + 1)/3 = 3.0]), suggesting that the instrument can distinguish respondents into low, medium, and high literacy levels. Item reliability of 0.66 with an item separation of 1.65 yields about two difficulty strata ([H = 2.5]), indicating moderate differentiation among item difficulty levels. The significant chi-square value (χ^2 = 597.12, df = 296, $p < .01$) reflects meaningful variability among respondents' abilities.

Table 1. Reliability and Separation

	Value
Person Reliability	0.71
Person Separation	1.98
Item Reliability	0.66
Item Separation	1.65
Cronbach alpha	0.74
Chi-square	597.12** with 296

Item Fit Statistics

Item fit statistics in the Rasch model are used to determine whether each item functions as expected in measuring a single latent trait, which in this study is chemical literacy among preservice chemistry teachers through ethnochemistry-integrated two-tier multiple-choice questions. Fit is commonly evaluated using Infit and Outfit mean square (MNSQ) values and their standardized forms (ZSTD). Infit is more sensitive to unexpected responses close to a respondent's ability level, whereas Outfit is influenced by outliers and unexpected responses far from the ability estimate. Generally, MNSQ

values between 0.5–1.5 and ZSTD values within ±2 indicate acceptable fit. Positive point-measure correlations (Pt Mea Corr) further confirm that items align with the intended construct (Wang & Lin, 2020; Aziz, et al., 2022).

Based on Table 2, most items demonstrate satisfactory fit to the Rasch model. Items 1, 4, 6, 7, 8, 9, and 10 show slightly low MNSQ values (<1.0), suggesting more predictable response patterns, but still within acceptable limits. Items 2 and 3 show higher MNSQ values (Infit MNSQ=1.15 and 1.30), indicating more variability, yet still acceptable. Item 5 shows the highest misfit tendency (Infit MNSQ=1.43; Outfit ZSTD=2.1), implying some unexpected response behavior and potential construct-irrelevant variance, though its Pt Mea Corr (0.53) remains positive and adequate (Smith, 2019). Pt Mea Corr values across items (0.53–0.61) indicate good discrimination and support item validity.

Table 2. Item Fit Statistics

No	Item	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Pt Mea Corr
1	Item 1	0.89	-0.5	0.88	-0.5	0.56
2	Item 2	1.15	0.8	1.08	0.4	0.55
3	Item 3	1.3	1.4	1.2	0.9	0.53
4	Item 4	0.85	-0.6	0.86	-0.4	0.56
5	Item 5	1.43	1.8	1.42	2.1	0.53
6	Item 6	0.88	-0.6	0.84	-0.6	0.55
7	Item 7	0.96	-0.2	0.91	-0.3	0.56
8	Item 8	0.92	-0.3	0.95	-0.1	0.61
9	Item 9	0.79	-1.0	0.79	-0.9	0.6
10	Item 10	0.78	-1.1	0.73	-1.2	0.6

Unidimensionality

Unidimensionality in the Rasch model refers to the assumption that an instrument measures a single dominant latent trait, ensuring that item responses are primarily driven by one construct. This assumption is commonly evaluated using principal component analysis (PCA) of standardized residuals, where the proportion of variance explained by Rasch measures indicates the strength of the intended dimension. As shown in Table 3, the total raw variance in observations is 14.1 eigenvalue units (100%), and the raw variance explained by measures is 4.1 eigenvalue units, accounting for 40.9% of the empirical variance (modeled 40.5%). This percentage exceeds the commonly suggested minimum criterion (≥20%), indicating that the instrument has a strong primary dimension. Additionally, variance contributions from persons (16.4%) and items (24.5%) suggest that both respondent ability and item difficulty contribute meaningfully to the measurement structure, supporting the intended unidimensional construct.

Table 3. Standrardized Residula Variance

	Empirical (Eigenvalue)	Empirical (%)	Modeled (%)
Total raw variance in observations	14.1	100.0%	100.0%
Raw variance explained by measures	4.1	40.9%	40.5%
Raw variance explained by persons	1.5	16.4%	16.3%
Raw variance explained by items	2.6	24.5%	24.2%

Discussion

The Rasch analysis results indicate that the ethnochemistry-integrated two-tier multiple-choice instrument functions effectively in mapping preservice chemistry teachers' chemical literacy across varying competence levels. The Wright map shows that most respondents are located around the mean ability range (−1 to +1 logits), suggesting that participants generally possess moderate chemical literacy, with fewer individuals at extreme ability levels. This pattern implies that while many preservice teachers can interpret chemical concepts embedded in cultural practices, there remains variability in their ability to connect ethnochemical contexts with deeper scientific reasoning. This finding aligns with the literature indicating that chemical literacy in Indonesia remains relatively low, partly because the socio-cultural dimension has not been optimally integrated into chemistry learning despite the close connections between many

chemistry concepts and everyday cultural practices, and learning outcomes may vary due to demographic factors that have still been rarely examined systematically in chemical literacy research (Ashari & Munawwarah, 2025).

The item difficulty distribution is also well spread, indicating that the test provides appropriate challenges across the ability continuum. Items with higher logit positions (e.g., Item 8 and Item 4) demand higher-order reasoning such as quantitative analysis and equilibrium interpretation within cultural contexts, whereas easier items (e.g., Item 1 and Item 6) assess more foundational chemical understanding linked to familiar practices.

In terms of measurement quality, reliability and separation indices support the instrument's adequacy for research and educational evaluation. Person reliability (0.71) and Cronbach's alpha (0.74) demonstrate acceptable internal consistency, indicating that the instrument consistently differentiates respondents' chemical literacy levels.

Item reliability (0.66) and item separation (1.65) indicate moderate stability of item difficulty estimates, implying that the item hierarchy is reasonably consistent but could be strengthened by additional items or larger samples. Item fit statistics confirm that most items align with Rasch expectations, with acceptable Infit/Outfit values and positive Pt Mea Corr, supporting construct validity. Prior Rasch studies emphasize that acceptable Infit/Outfit values combined with positive point-measure correlations are typically taken as evidence that items contribute cohorently to the intended latent construct (Dwi, et al., 2021).

However, Item 5 shows borderline misfit (Outfit ZSTD=2.1), indicating potential unexpected response patterns that may require revision in wording, cultural context clarity, or distractor functioning. Unidimensionality evidence is strong, with 40.9% variance explained by measures, reinforcing that the instrument primarily measures a single construct of chemical literacy.

This study offers important implications for chemistry education by providing a validated ethnochemistry-integrated two-tier instrument to measure chemical literacy among pre-service chemistry teachers through culturally contextualized chemistry content. The instrument supports assessment of chemical literacy beyond factual recall by emphasizing conceptual interpretation and scientific reasoning when responding to real-life cultural practices. Embedding ethnochemistry promotes culturally responsive assessment, helping teacher education programs connect chemistry concepts with local knowledge systems and everyday phenomena relevant to West Nusa Tenggara.

The use of Rasch model analysis strengthens the measurement quality by enabling item calibration and person ability estimation on an interval logit scale, supporting objective interpretation of chemical literacy levels. Rasch measurement has been widely recommended in educational assessment because its logit scaling provides interval-level estimates that allow more defensible comparisons of learner ability than raw scores alone (Adi, et al., 2022). The validated item set can assist lecturers and researchers in evaluating literacy-oriented learning outcomes, refining instruction, and developing future ethnochemistry-based assessments aligned with chemical literacy indicators.

Several limitations should be considered in interpreting this study.

First, the participants consisted of only 35 pre-service chemistry teachers selected through convenience sampling from a single public university in West Nusa Tenggara, limiting the representativeness of the sample and reducing the generalizability of the findings to other institutions or regions.

Second, the ethnochemistry contexts embedded in the items were derived from local cultural practices, which may not be equally familiar to respondents from different cultural backgrounds, potentially affecting item interpretation and response consistency.

Third, the study used a paper-based, face-to-face administration format; therefore, the instrument's functioning may differ in online or computer-based testing environments.

Fourth, although expert judgment was used to validate content relevance and cultural appropriateness, expert feedback is inherently subjective and may vary depending on the experts' experiences and perspectives regarding chemical literacy and ethnochemistry integration.

CONCLUSION

This study developed and validated an ethnochemistry-integrated two-tier multiple-

choice instrument designed to assess chemical literacy among preservice chemistry teachers.

The integration of ethnochemistry enabled chemical literacy assessment to be grounded in local cultural practices, supporting contextual interpretation of chemistry concepts and reasoning-based responses. The two-tier structure strengthened the measurement of chemical literacy by combining content understanding with justification selection, ensuring that responses reflect both conceptual and reasoning dimensions of literacy.

Rasch model analysis provided robust evidence of instrument quality by evaluating item functioning, measurement consistency, and construct representation through item calibration, fit analysis, reliability indices, separation measures, and unidimensionality testing. This study provides a reliable, culturally responsive diagnostic tool that can support teacher education programs in identifying preservice teachers' strengths and weaknesses in chemical lietracy and informing evidence-based instructional design.

The results support the use of this instrument as a culturally responsive tool for chemical literacy assessment in teacher education contexts, particularly for preservice chemistry teachers in West Nusa Tenggara. The novelty of this work lies in the integration of ethnochemistry with a two-tier assessment format validated through the Rasch model, producing a calibrated item hierachy that enables more precise and fair measurement of chemical literacy.

RECOMMENDATION

Future research should involve larger and more diverse samples across multiple universities and regions to improve generalizability and strengthen the stability of item calibration within the Rasch model framework. Further refinement of items is recommended, particularly for items showing borderline fit patterns, by revising wording, improving clarity of ethnochemical contexts, and enhancing the quality of reasoning options. Additional ethnochemistry contexts from other Indonesian regions can be incorporated to broaden cultural representation and ensure the instrument remains adaptable across different local settings.

Researchers may also explore digital or computer-based administration to evaluate whether the instrument functions consistently in online testing environments. In teacher education practice, the instrument can be applied to diagnose chemical literacy profiles and support the design of literacy-oriented chemistry instruction grounded in local wisdom and contextual chemistry learning.

BIBLIOGRAPHY

Adam, W., Qudratuddarsi, H., Ningthias, D. P., Rahmadhani, A., & Noviana, E. (2025). Validation of Pre-Service Science Teacher Artificial Intelligence Competence Self-Efficacy (AICS): Rasch Model Analysis. *Jurnal Ilmiah Profesi Pendidikan*, *10*(2), 1985-1995.

Adi, N. R. M., Amaruddin, H., Adi, H. M. M., A'yun, I. L. Q. (2022). Validity and Reliability Analysis Using the Rasch Model to Measure the Quality of Mathematics Test Items of Vocational High Schools. *Journal of Educational Research and Evaluation*, 11(1), 103-113.

Alnahdi, A. H., Alsubiheen, A. M., & Aldaihan, M. M. (2025). Rasch Measurement Model Supports the Unidimensionality and Internal Structure of the Arabic Oswestry Disability Index. *Journal of Clinical Medicine, 14*(4), 1259.

Alwathoni, M., Saputro, S., Yamtinah, S., Masykuri, M., & Suwahono, S. (2024). MOOC-Based PBL Model to Improve Chemical Literacy Skills In Content, P rocedural and Epistemic Knowledge Aspects. *International Journal of Pedagogy and Teacher Education*, *8*(2), 253-268.

Ashari, A., & Munawwarah (2025). Ethnochemistry Supports 21st Century Skills: Systematic Literature Review. *Hydrogen: Jurnal Kependidikan Kimia*, *13*(5), 1044-1049.

Aziz, A. A., Rahman, N. A., & Khalid, F. (2022). Rasch analysis of an educational assessment instrument: Evidence from item fit and person–item targeting. *Education Sciences*, 12(9), 612.

Bond, T. G., & Fox, C.M. (2021). Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Routledge.

Boone, W. J., Staver, J., & Yale, M. (2021). Rasch analysis in the human sciences. *Springer*.

Dewi, C. A., Yahdi, Y., & Sanova, A. (2024). Ethnochemistry-Based E-Module: Does it Effect on Improving Students' Chemical Literacy. *Journal of Innovation in Educational and Cultural Research*, *5*(4), 568-577.

Dwi, A., Yamtinah, S., Mahardiani, L., & Sauptro, S. (2021). A Rasch Analysis of Item Quality of the Chemical Literacy Assessment for Investigasting Student's Chemical Literacy on Chemical Rate Concepts. European Journal of Educational Research, 10(4), 1769-1779.

Fischer, H. E., Boone, W. J., & Neumann, K. (2023). Quantitative research designs and approaches.

In *Handbook of research on science education* (pp. 28-59). Routledge.

Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. *International Journal of Education & Language Studies*, *1*(2), 72-77.

Hany, N. C., & Purwaningsih, D. (2025). Collaborative Learning Strategies in Improving High School Students' Chemistry Literacy: A Systematic Literature Review. *Jurnal Pendidikan Kimia Indonesia*, *9*(1), 42-52.

Hasanah, S. M., Rahayu, S., Sulistina, O., Alsulami, N. M. (2025). Creating and Validating a Chemical Literacy Instrument Aligned Sustainability Contexts: Construct and Rasch Measurement Analysis. *Indonesian Journal of Science Education,* 13(2), 408-425.

Hidayat, R., Qudratuddarsi, H., Ayub, A. F. M., & Latif, I. N. A. (2025). Psychometric Properties Of The Perma-Profiler For Indonesian College Students: A Rasch Modelling Analysis. *Journal of Institutional Research South East Asia*, *23*(1).

Hidayat, R., Qudratuddarsi, H., Mazlan, N. H., & Zeki, M. Z. M. (2021). Evaluation of a test measuring mathematical modelling competency for Indonesian college students. *Journal of Nusantara Studies (JONUS)*, *6*(2), 133-155.

Ismail, M. S., Din, M. H., & Jusoh, M. S. (2021, July). Predictive modelling using Rasch's person-item map: Intrepreting and assessing in Malaysia manufacturing firms. In *AIP Conference Proceedings* (Vol. 2347, No. 1, p. 020288). AIP Publishing LLC.Jamieson, M. K., Govaart, G. H., & Pownall, M. (2023). Reflexivity in quantitative research: A rationale and beginner's guide. *Social and Personality Psychology Compass*, *17*(4), e12735.

Jumriani, J., Qudratuddarsi, H., Rahmah, N., Ningthias, D. P., & Indriyanti, N. (2025). Validation Of Instrument to Measure Gen-Z Pre-Service Teacher Chemistry Virtual Lab Acceptance and Use. *Chemistry Education Practice*, *8*(1), 128-136.

Junaidi, E., Sudatha, I. G. W., Suartama, I. K., & Santosa, M. H. (2025). Ethnochemistry In Chemistry Learning: Insights from Indonesian Local Wisdom. *Jurnal Pendidikan MIPA*, *26*(3), 1642-1658.

Kassiavera, S., Suparmi, A., Cari, C., & Sukarmin, S. (2024). Application of Rasch model in two-tier test for assessing students' reasoning in physics education. *Journal of Baltic Science Education*, 23(6), 1228–1245.

Lamb, K. N., Boedeker, P., & Kettler, T. (2025). Measuring creative self-efficacy: Instrument development and validation. *Thinking Skills and Creativity*, *56*, 101738.

Manik, A. (2024). Analysis of Science Literacy and Chemical Literacy Content on Green Chemistry Materials in Chemistry Textbooks Class X High School in Palangka Raya City. *Chemistry in Education*, *13*(1), 31-41.

Medvedev, O. N., & Krägeloh, C. U. (2025). Rasch measurement model. In *Handbook of assessment in mindfulness research* (pp. 131-147). Cham: Springer Nature Switzerland.

Marfuatun, M., Nahadi, N., Yuliani, G., & Hernani, H. (2024). Assessing Chemical Literacy of Pre-service Chemistry Teacher: Rasch Analysis. In *SHS Web of Conferences* (Vol. 205, p. 06001). EDP Sciences.

Murashko, T., Kolesovs, A., & Ruza, A. (2025). Imagining the Future: Instrument Development and Initial Validation. *World Futures*, 1-21.

Nash, R., & Bradley, S. (2021). Using Rasch analysis to evaluate science assessment instruments. *Research in Science Education*, 51(6), 1521–1543.

Natania, A., Subarkah, C. Z., & Sundari, C. D. D. (2025). Ethnochemistry of Kampung Naga: Local Wisdom as Green Chemistry Learning in E-Book Form. *Orbital: Jurnal Pendidikan Kimia*, *9*(2), 174-188.

Ningroom, R. A. A, N., Yamtinah, S., & Riyadi. (2025) A two-tier multiple choice diagnostic test to find student misconceptions about the change of matter. *Journal of Education and Learning*, 19(2), 1144-1156.

Ningtyas, A. I., Yamtinah, S., Saputro, S., Shidiq, A. S., & Rahayu, S. (2025). Content Validity Analysis of a Virtual Reality Based Two-Tier Multiple Choice Assessment Instrument with Ethnochemistry to Early Detect Misconceptions in Reaction Rate Topics. *Jurnal Penelitian Pendidikan IPA*, *11*(5), 960-969.

Nugraheni, A. R. E., & Srisawasdi, N. (2025). Development of pre-service chemistry teachers' knowledge of technological integration in inquiry-based learning to promote chemistry core competencies. *Chemistry Education Research and Practice*, *26*(2), 398-419.

Pamungkas, Y. P., Yamtinah, S., Saputro, S., & Shidiq, A. S. (2024). Detection Student Misconception in Chemical Bonding using a Virtual Reality-Integrated Two-Tier Multiple-Choice Instrument with Ethnochemistry Context. *Jurnal Pendidikan MIPA*, *25*(3), 1498-1517.

Pebriani, N. I., & Aini, F. Q. (2025). The Relationship Between Prior Knowledge and Students' Chemical Literacy. *Jurnal Pijar Mipa*, *20*(5), 970-975.

Peeters, M. J., & Augustine, J. M. (2023). Using Rasch measurement for instrument rating scale refinement. *Currents in Pharmacy Teaching and Learning*, 15(2), 110–118.

Post, M. W. M., Fellinghauer, C. S., Charlifue, S., New, P. W., Forchheimer, M. B., & Tate, D. G. (2022). Rasch analysis of the International Quality of Life Basic Data Set Version 2.0. *Archives of Physical Medicine and Rehabilitation*, 103(11), 2120–2130.

Pratama, F. I., Joronavalona, R., Annisa, D., Saputri, F., Afikah, A., Wijayanti, P., ... & Naqsyahbandi, F. (2025, September). Differences in chemical literacy of chemistry education students based on study

year level. In *AIP Conference Proceedings* (Vol. 3354, No. 1, p. 040027). AIP Publishing LLC.

Qudratuddarsi, H., Meivawati, E., & Saputra, R. (2024). Pelatihan Penelitian Metode Kuantitatif dan Systematic Literature Review bagi Dosen dan Mahasiswa. *Beru'-beru': Jurnal Pengabdian kepada Masyarakat*, *3*(1), 22-32.

Qudratuddarsi, H., Rahmah, N., & Indriyanti, N. (2025). Analysis of the attitudes of generation z prospective science teachers toward the chemistry virtual laboratory. *Arfak Chem: Chemistry Education Journal*, *8*(1), 651-665.

Schoonenboom, J. (2023). The fundamental difference between qualitative and quantitative data in mixed methods research. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* (Vol. 24, No. 1). DEU.

Sjöström, J., Yavuzkaya, M., Guerrero, G., & Eilks, I. (2024). Critical chemical literacy as a main goal of chemistry education aiming for climate empowerment and agency. *Journal of Chemical Education*, *101*(10), 4189-4195.

Smith, R. M. (2019). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, 20(3), 287–300.

Sutrisno, Manuharawati, & Masriyah. (2025, November). Assessing beliefs about statistics of prospective mathematics teachers using the rasch wright map. In *AIP Conference Proceedings* (Vol. 3333, No. 1, p. 020007). AIP Publishing LLC.

Tennant, A., & Conaghan, P. (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.

Wang, W., & Lin, Y. (2020). Rasch validation of an assessment instrument: Item fit, reliability, and unidimensionality. *Frontiers in Psychology*, 11, 569.

Wardani, S. F., Yamtinah, S., Mulyani, B., Susilowati, E., Ulfa, M., Masykuri, M., & Shidiq, A. S. (2024). Differentiated learning: analysis of students' chemical literacy on chemical bonding material through culturally responsive teaching approach integrated with ethnochemistry. *Jurnal Penelitian Pendidikan IPA*, *10*(4), 1747-1759.

Yuendita, D., & Rohaeti, E. (2025). Research Trends on the Integration of Ethnoscience in the Learning of Chemistry as the Development of Chemical Literacy: A Systematic Review. *Indonesian Journal of Educational Research and Review*, *8*(1), 210-222.