Jurnal Pendidikan Kimia (JPKIM)
p-ISSN: 2085-3653| e-ISSN: 2549-3116
https://jurnal.unimed.ac.id/2012/index.php/jpk/index

OPEN ACCESS
10.24114/jpkim.v16i3.67118
J. Pendidik. Kim 2024, 16(3), 251– 256

Original Research Article

AI-based assessment: Aiken index analysis of content validity in scientific literacy instruments on the basic laws of chemistry with ethnochemistry

Fila Delfia Zai
Department of Chemistry Education, Universitas Sebelas Maret, Surakarta 57126, Indonesia

Sri Yamtinah - *https://orcid.org/0000-0001-5616-4897*
Department of Chemistry Education, Universitas Sebelas Maret, Surakarta 57126, Indonesia

Ari Syahidul Shidiq - *https://orcid.org/0000-0003-4388-5681*
Department of Chemistry Education, Universitas Sebelas Maret, Surakarta 57126, Indonesia

Hayuni Retno Widarti - *http://orcid.org/0000-0001-5573-9345*
Department of Science Education, Universitas Negeri malang, Malang 65145, Indonesia

Antuni Wiyarsi - *http://orcid.org/0000-0001-5573-9345*
Department of Science Education, Universitas Negeri Yogyakarta, Yogyakarta 55281, Indonesia

Corresponding authors: Sri Yamtinah (e-Mail: jengtina@staff.uns.ac.id)

Citation: Zai, F., Yamtinah, S., Shidiq, A.S., Widarti, H., & Wiyarsi, A. (2024). AI-based assessment: Aiken index analysis of content validity in scientific literacy instruments on the basic laws of chemistry with ethnochemistry. Jurnal Pendidikan Kimia (JPKIM), 16(3), 251 – 256. https://doi.org/10.24114/jpkim.v16i3.67118

ARTICLE INFO

Keywords:
Aiken index;
Artificial intelligence;
Content validity;
Science literacy assessment

History:
♦Received - 20 Dec 2024
♦Revised - 24 Dec 2024
♦Accepted - 26 Dec 2024

ABSTRACT

Scientific Literacy Assessment (SLA) is used to measure science literacy on basic chemical law material that has been qualified and tested. This study aims to prove the content validity of ethnochemistry SLA instruments integrated with artificial intelligence (AI) systems using Aiken's validity index so that they can measure students' scientific literacy in basic chemical law material. The research design uses the ADDIE development model (Analysis, design, development, implementation and evaluation which is simplified into analysis, design, and development. Content validity data was obtained from 10 experts, namely 7 chemistry education lecturers from UNS, UNY, and UM and 3 chemistry teachers in Surakarta through focus group discussions (FGD). The content validation test is based on aspects of context, language, and construct. A valid instrument is indicated by an Aiken index value of more than or equal to 0.73. The AI-integrated ethnochemistry SLA instrument on basic laws of chemistry material produced 15 valid items with an Aiken index greater than 0.73.

Introduction

In the 21st century, the education system must continue to evolve to keep pace with modern times. Dragos & Mih (2015) emphasized that science literacy is crucial in helping individuals understand their environment. To address the challenges posed by globalization and technological advancements, science education in schools needs to adapt to societal changes. A strong understanding of science is considered the foundation of students' science literacy (Zou et al., 2024). Furthermore, UNESCO (2021) highlights the importance of science as a tool for enhancing the quality and accessibility of scientific findings and processes. Therefore, science education facilitates knowledge transfer and contributes significantly to improving students' science literacy.

Science comprises three fundamental interconnected components: product, scientific process, and scientific attitude. Science as a product encompasses facts, concepts, procedures, principles, and laws of nature. Science as a process signifies that scientific knowledge is obtained from a scientific process or scientific activity. Science as an attitude refers to the scientific attitude that serves as the foundation for the scientific process to produce scientific products. These three components are instrumental in evaluating the fundamental science knowledge possessed by students (DinanThompson et al., 2015; Goss et al., 2022).

Science literacy is defined as the capacity to comprehend and engage in discourse concerning scientific subjects, in addition to the aptitude for analytical and logical thinking (Aulia et al., 2018). The Program for International Student Assessment (PISA) delineates science literacy as the foundation of the capacity to engage with scientific subjects and concepts, encompassing the aptitude to elucidate phenomena through a scientific lens, evaluate and devise scientific inquiries, and interpret data and evidence scientifically (DinanThompson et al., 2015; Goss et al., 2022). The capacity of students to cultivate an understanding of the conceptual framework of the discipline and to interweave it with a comprehensive science understanding is designated as conceptual science literacy (Schwartz et al., 2006). According to Shwartz et al. (2006), indicators of science literacy encompass content aspects, context aspects, high-order learning skills

(HOLS), and attitudinal aspects. Crucial aspects of scientific literacy, such as scientific knowledge, comprehension of fundamental concepts, and the capacity to apply scientific knowledge in everyday scenarios, can be adequately measured.

The results of the PISA in 2015 indicate that the science literacy of Indonesian students is ranked 62nd out of 72 participating countries, with a score of 403, which is considerably below the OECD average of 500. In 2018, Indonesia's science literacy score decreased to 396, indicating a persistent challenge in enhancing students' science literacy skills. Notably, approximately 70% of 15-year-olds demonstrate reading literacy levels that fall below the minimum competency standards. These students demonstrate an ability to recognize routine information from short texts and to comprehend simple procedures. The PISA 2022 results indicate an improvement in Indonesia's science literacy ranking, with an increase of six positions. However, it is noteworthy that the international literacy average declined by 18 points, and Indonesia's score decreased by 12 points (Kemendikbudristek, 2023). In the context of science application, Indonesian students frequently encounter challenges in establishing connections between science knowledge and real-life events or phenomena. Furthermore, the majority of Indonesian students demonstrate a limited capacity for logical, systematic, and rational thinking (Permanasari, 2016). This observation underscores the necessity of integrating scientific practices into the educational curriculum to foster students' comprehension of scientific concepts (Berland et al., 2016).

The dearth of science literacy in Indonesia can be attributed to the nation's educational curriculum, pedagogical methodologies, and evaluation systems, which are not adequately aligned to foster the attainment of science literacy (Rubini et al., 2018; Jufri et al., 2019; Asikin et al., 2019). Assessment, defined as evaluation, constitutes a pivotal component within educational practices and accountability systems in the educational process (DinanThompson et al., 2015; Goss et al., 2022). In order to enhance students' science literacy, it is imperative to develop assessment instruments that are meticulously designed to evaluate science literacy, thereby ensuring that students are well-versed in the diverse facets of the scientific realm (Rusilowati et al., 2018; Muniroh et al., 2022; Chasanah et al., 2022). An analysis of observations and interviews with chemistry teachers at State Senior High School (SMAN) 6 Surakarta, SMAN 1 Banyudono, and SMAN 1 Teras Boyolali reveals a limitation of existing chemistry subject assessment instruments in summative assessments. These instruments merely measure students' comprehension of the material, neglecting to evaluate their science literacy skills. Furthermore, the questions posed in these assessments lack narratives that are pertinent to real-life applications, despite the frequent relevance of the taught material to real-world contexts. Consequently, there is a necessity for the development of a Science Literacy Assessment (SLA) to evaluate students' science literacy skills.

The assessment instrument can take the form of a description test, which has been demonstrated to be an effective method for measuring or assessing the results of a complex learning process (Putri et al., 2022). Descriptive tests have been shown to facilitate the evaluation of students' ability to apply science in a practical context and to measure students' thinking skills, which are critical components of science literacy. However, it should be noted that descriptive questions are not without their limitations. These limitations include a lack of objectivity and efficiency in assessment and the difficulty for teachers to provide appropriate feedback (Rosyidi, 2020). A study of interviews conducted with chemistry teachers at SMAN 6 Surakarta, SMAN 1 Banyudono, and SMAN 1 Teras Boyolali revealed that descriptive questions necessitate greater time and accuracy to evaluate answers individually. This approach is suboptimal in terms of efficiency due to the time-consuming nature of the process.

By surmounting these challenges, the integration of artificial intelligence (AI) technology into the assessment process becomes a viable prospect. AI possesses considerable potential to support educators in administering assessments that are expeditious, objective, and adaptable. Specifically, AI can generate adaptive assessments that are capable of adjusting to student responses in real-time, providing immediate feedback, and enhancing the accuracy of the assessment process (Khlaif et al., 2024). Furthermore, AI can assist in identifying student weaknesses by analyzing answer patterns and providing targeted feedback to facilitate improvement (Dwivedi, 2021).

To enhance students' science literacy, it is imperative to optimize the chemistry learning process in educational institutions. However, chemistry education has historically prioritized scientific principles, potentially leading to a disconnect between students and the environment, technology, and society (Morales-Doyle et al., 2019). Additionally, chemistry involves the use of submicroscopic models to elucidate various phenomena (Laohapornchaipan & Chenprakhon, 2024). The subject matter of chemistry encompasses not only a body of facts and principles but also methods for deriving these concepts and fostering a scientific attitude. The inherent complexity and interconnectedness of chemistry concepts contribute to their abstract nature and the challenges in comprehending them, particularly within educational settings. Consequently, a comprehensive understanding of chemistry necessitates an integration of theoretical processes and real-life applications (Cooper & Stowe, 2018).

Culture-based chemistry learning underscores the significance of cultural competence among educators, encompassing the comprehension and respect for students' cultural heritages. The incorporation of cultural values into chemistry education can be facilitated through the conceptual framework of ethnochemistry (Rahmawati et al., 2019). This field of study encompasses the examination of the interplay between culture and chemistry, with a particular focus on the manner in which cultural beliefs, practices, and knowledge affect chemical concepts and processes (Wardani et al., 2023). In a related study, Pebrianti et al. (2024) developed an acid-base module derived from ethnochemistry problems, employing the Aiken formula to ascertain its validity. This module has been shown to enhance students' science literacy skills. Concurrently, Asda et al. (2023) undertook a study that led to the formulation of ethnochemistry charged buffer solution worksheets, which were found to be valid and efficacious in enhancing students' science literacy skills.

The fundamental principles of chemistry are intimately intertwined with day-to-day life, as evidenced by the combustion of paper in an open space. The foundational laws of chemistry encompass the law of conservation of mass, the law of multiple comparisons, the law of permanent comparison, the law of volume comparison, and the Avogadro hypothesis. The interconnected nature of these principles underscores the necessity for a comprehensive understanding of each concept, as a lack of clarity on one concept can impede the grasp of others (Wasonowati, 2014). For instance, misconceptions regarding the law of conservation of mass persists, with a significant proportion of students erroneously perceiving reactions as the

cause of alterations in a substance's mass (Putri et al., 2022). Integrating ethnochemistry with fundamental chemical law materials that are pertinent to everyday life has the potential to enhance students' science literacy in chemical materials.

The objective of this research is to develop a Science Literacy Assessment (SLA) instrument that incorporates ethnochemistry content and an Artificial Intelligence (AI) system, focusing on the fundamental principles of chemistry. The instrument's validity will be substantiated by the Aiken index, with considerations given to three aspects: context, language, and construct.

Methods

This research employs the Research and Development (R&D) method, which is a systematic approach that combines research and practical applications to innovate and improve products or processes. The development research design utilizes the ADDIE development model, which is simplified into three stages: analysis, design, and development. The analysis stage includes a literature study, observation, interviews, and analyzing the software to be used. The design stage involves identifying the subject matter, including ethnochemistry, designing question indicators, and preparing a validation questionnaire. The development stage encompasses creating SLA questions, making answer keys and prompts, conducting product validation, product revision, and individual limited trials.

The analysis of the instrument validation results was conducted qualitatively to assess the ethnochemistry and AI-based SLA instruments that were developed.The qualitative data was derived from the validator's comments and input regarding the content, language, and presentation of the instrument.The content validity of the instrument was evaluated using the Aiken index. This index was determined through expert agreement via focus group discussion (FGD), wherein ten validators assigned scores of 1-4 to each question item based on content, language, and construct.The validity of a question is deemed satisfactory if the V Aiken index is greater than or equal to 0.73 for a total of 10 raters (Aiken, 1980).

$V = \Sigma S / [n(C-1)]$ with $S = R-Lo$

Description:

V : Aiken's validity index
S : rater's score minus the lowest score in the category
R : score given by the evaluator
Lo : lowest rating score (1)
C : highest rater score (4)
N : number of validators (raters)

Results and Discussion

The content validity of the AI-based ethnochemical SLA instrument on fundamental chemical law material that has been developed is established through expert consensus, as an instrument, both test and non-test, is deemed valid if experts concur that the instrument possesses the capacity to measure the intended ability.The content validity assessment of the SLA instrument that has been developed is directed towards three domains: (1) content aspects, (2) language aspects, and (3) construct aspects. Table 1 presents the assessment grid of the AI-integrated ethnochemical SLA instrument. The results of the Aiken index analysis obtained from the focus group discussion (FGD) were obtained from 10 raters through a validity sheet containing a rating scale from 1-4 on the SLA instrument that had been developed. Subsequently, the validation of the assessment results by the validator was calculated using the Aiken index, as illustrated in Table 2.

Table 1. Presents an assessment grid for an AI-based ethnochemistry SLA instrument on the basic laws chemistry.

Aspect	Assessment grids
Aspect of content	The material being taught is pertinent to the subject matter.
	The concepts or information presented in the question items are accurate and suitable, and they align with the fundamental competencies.
	There are no factual errors in the question items.
Aspect of language	The text demonstrates a commendable level of grammatical proficiency, exhibiting proper sentence structure that is conducive to effective student comprehension.
	The absence of spelling, punctuation, or word usage errors in the question items further attests to the text's linguistic accuracy and precision.
	The inquiries have been meticulously arranged in a coherent sequence, with each statement being intricately linked to its subsequent counterpart, thereby establishing a unified problem-solving framework.
	The questions are not repetitive or overlapping, which contributes to a logical progression in the inquiry process.
Aspect of construct	The instructions are meticulously articulated, ensuring optimal clarity and accessibility.

The degree of content validity is contingent upon the alignment between the question items and the designated indicators. This study utilizes a four-tiered response scale to ascertain the validity of responses, categorizing them as follows: irrelevant (TR) score 1, less relevant (KR) score 2, moderately relevant (CR) score 3, and relevant (R) score 4. The validity of a question is deemed adequate if the Aiken's V index exceeds or equals 0.73. The Aiken's validity index is classified into three categories: low, medium, and high (Aiken, 1985). Scores falling below 0.4 are designated as low, scores ranging from 0.4 to 0.8 are classified as medium, and scores exceeding 0.8 are categorized as high. When the Aiken index value approaches 1.0, the item is considered to be of higher quality (Ulfah et al., 2020). The validation process, conducted by ten assessors/validators, resulted in a V Aiken index value greater than 0.73 (Vtabel). The calculation of Aiken's index on 15 items of SLA instrument questions that have been developed obtained 1 item with moderate validity, and 14 items with high validity. This shows that the AI-based ethnochemistry SLA instrument on basic chemical law material that has been

developed is valid and can be used for students in the learning process. The SLA instrument developed has fulfilled the aspects of content, language, and construct.

Table 2. Presents the Aiken Index analysis results on the AI-based ethnochemistry SLA instrument on the basic laws of chemistry.

Question Item	V Value	V Table	Conclusion
1	0.90	0.73	Valid
2	0.80	0.73	Valid
3	0.86	0.73	Valid
4	0.90	0.73	Valid
5	0.88	0.73	Valid
6	0.96	0.73	Valid
7	0.91	0.73	Valid
8	0,90	0.73	Valid
9	0.91	0.73	Valid
10	0.90	0.73	Valid
11	0.96	0.73	Valid
12	0.93	0.73	Valid
13	0.95	0.73	Valid
14	0.97	0.73	Valid
15	0.93	0.73	Valid

The results of the FGD analysis of the SLA instrument obtained suggestions and input from the assessor on the quality of the AI-based ethnochemistry SLA instrument on basic chemical law material contained in the instrument validity sheet. Suggestions and input from validators are presented in Table 3.

Table 3. Presents the outcomes of the FGD analysis of AI-based ethnochemistry SLA instruments on the basic laws of chemistry

Suggestions and feedback	Researcher follow-up
Some narratives are less helpful in directing students to answer questions	The narrative must be revised to achieve clarity, relevance, and focus on information that supports the question
Review the narrative used	Reviewing the narrative with the involvement of validators to ensure clarity and conformity with the competencies measured.
Added some chemical reactions required in the reading	Adding appropriate reaction examples to reinforce context and support student understanding
It is imperative to make comparisons for all elements in the law of multiple comparisons sub-matter.	Complement the reading with relevant tables or data on elemental mass ratio to clarify concepts
Improve grammar and question constructs	Reviewing questions to ensure appropriate grammar and clear question structure
A review of the informative and reasonable stimulus provided with each reading material is still necessary	The stimulus should be revised by the addition of data or information that supports the question.
The use of conjunctions at the start of sentences should be avoided.	The conjunction is relocated to its proper position, thereby enhancing the sentence.
The construction and language of the questions should be improved.	Use straightforward, simple sentences, and replace terms that are difficult for students to understand

Table 4. AI Integrated SLA prompt format

Sub-content	Description
Role and objective	Contains an overview of what character the AI should be to evaluate student answer
Question	Contains questions
Answer key	Contains answer key
Scoring rubric	Contains details of scoring guidelines. As for this system, the maximum score for each question has the same value to make it easier for the AI system to work.
Evaluation instructions	Contains details on how the AI will evaluate student answers.
Example feedback structure	Contains an example of the structure of the feedback that students get. Here is the format: Total Score: [Insert Total Score Here] out of 10 points Scoring Breakdown: Question : [Insert Score] - [Reason for Score] Feedback: "..."
Final comment	It summarizes how the student performed along with highlighting strengths and areas for improvement.

Following the revision of the AI-based ethnochemistry SLA instrument in response to expert input, the instrument will be integrated into ChemTest (cemytest.com), an automated assessment platform designed to assist chemistry teachers in evaluating student essay answers accurately and efficiently. The assessment is obtained from analyzing the AI, which is organized based on the prompt format. The prompt format is a method of organizing prompts that affect the results of the text. Additionally, the prompt format assists users in formulating requests (Dang et al., 2022). The AI-integrated SLA prompt format is presented in Table 4. The results of the content validity analysis of 15 items of the AI-based ethnochemistry SLA instrument on the basic laws of chemistry obtained a valid question instrument. However, of the 15 items, some changes are needed based on comments and suggestions from the people who checked the instrument to make it perfect so that it can be

used to measure students' science literacy on the basic laws chemistry. If the questions are improved, a small test will be done with 5 XI grade students in each of SMA Negeri 6 Surakarta, SMA Negeri 1 Teras Boyolali, and SMA Negeri 1 Banyudono.

Conclusion

The results of the development of the SLA instrument with ethnochemistry content integrated with the Artificial Intelligence (AI) system on basic chemical law material were declared valid. Based on the validity test results, this instrument has 14 items with a high level of validity and 1 item with a moderate level of validity, so it is feasible to use to support learning. However, even though it is considered feasible, this instrument still needs some improvements according to the input and suggestions of the validators so that it can function more optimally as an evaluation tool in measuring students' science literacy skills. The SLA instrument with ethnochemistry content integrated with the AI system was then used for a limited scale trial of 5 students in each of SMA Negeri 6 Surakarta, SMA Negeri 1 Teras Boyolali, and SMA Negeri 1 Banyudono.

Conflict of Interests

The author declares that there is no conflict of interest in this research and manuscript.

Acknowledgment

The authors express their gratitude to the Institute for Research and Community Service (LPPM) of Universitas Sebelas Maret, for supporting the Collaborative Research program between Universitas Sebelas Maret (UNS), Universitas Negeri Malang (UM), and Universitas Negeri Yogyakarta (UNY) under the Indonesia Collaboration Research (RKI) scheme with contract number 285/UN27.22/PT.01.03/2024. This support has significantly contributed to the successful implementation of this research.

References

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

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131-142. https://doi.org/10.1177/0013164485451012

Asda, V. D. (2023). Development of Buffer Solution Studentsâ€™ Worksheet Based on Problem Based Learning with Ethnochemistry to Improve Students Science Literacy Ability. *Jurnal Penelitian Pendidikan IPA*, 9(7), 5220-5227. https://doi.org/10.29303/jppipa.v9i7.4369

Asikin, N., & Yulita, I. (2019). Scientific literacy-based chemical teaching materials design of chemical solution materials on sea pollution context. *Jurnal Penelitian Pendidikan IPA*, 5(2), 204. https://doi.org/10.29303/jppipa.v5i2.249

Aulia, E. V., Poedjiastoeti, S., & Agustini, R. (2018). The effectiveness of guided inquiry-based learning material on students' science literacy skills. *Journal of Physics: Conference Series*, 947, p. 012049. https://doi.org/10.1088/1742-6596/947/1/012049

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, 53(7), 1082-1112. https://doi.org/10.1002/tea.21257

Chasanah, N., Widodo, W., & Suprapto, N. (2022). Pengembangan instrumen asesmen literasi sains untuk mendeskripsikan profil peserta didik. *PENDIPA Journal of Science Education*, 6(2), 474-483. https://doi.org/10.33369/pendipa.6.2.474-483

Chiu, M. (2022). My Journey in Chemistry Education. *Pure and Applied Chemistry*, 94(8), 919-941.

Cooper, M. M., & Stowe, R. L. (2018). Chemistry education research—From personal empiricism to evidence, theory, and informed practice. *Chemical Reviews*, 118(12), 6053-6087. https://doi.org/10.1021/acs.chemrev.8b00020

Dang, H., Mecke, L., Lehmann, F., Goller, S., & Buschek, D. (2022). How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models. arXiv preprint arXiv:2209.01390. https://doi.org/10.48550/arXiv.2209.01390

DinanThompson, M., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, 21(4), 485-503. https://doi.org/10.1177/1356336X15584087

Dragoş, V., & Mih, V. (2015). Scientific literacy in school. *Procedia-Social and Behavioral Sciences*, 209, 167-172. https://doi.org/10.1016/j.sbspro.2015.11.273

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Eilks, I., Witteck, T., & Pietzner, V. (2012). Eilks, I., Witteck, T., & Pietzner, V. (2012). The role and potential dangers of visualisation when learning about sub-microscopic explanations in chemistry education. *Ceps Journal*, 2(1), 125-145. https://doi.org/10.25656/01:5828

Goss, H. R., Shearer, C., Knowles, Z. R., Boddy, L. M., Durden-Myers, E. J., & Foweather, L. (2022). Stakeholder perceptions of physical literacy assessment in primary school children. *Physical Education and Sport Pedagogy*, 27(5), 515-530. https://doi.org/10.1080/17408989.2021.1911979

Jufri, A. W., Ramdani, A., Jamaluddin, J., & Azizah, A. (2019). Development of Scientific Literacy and Pedagogical Content Knowledge (PCK) of Prospective Science Teachers through Lesson Study-Based Courses. *Jurnal Penelitian Pendidikan IPA*, 5(2), 179-184. https://doi.org/10.29303/jppipa.v5i2.235

Kemendikbudristek. (2023, 5 Des). Peringkat Indonesia pada PISA 2022 Naik 5-6 Posisi Dibanding 2018. Diakses pada 14 Desember 2024, https://www.kemdikbud.go.id/main/blog/2023/12/peringkat-indonesia-pada-pisa-2022-naik-56-posisi-dibanding-2018

Khlaif, Z., Odeh, A., & Bsharat, T. R. K. (2024). Generative AI-Powered Adaptive Assessment. *Power of Persuasive Educational Technologies in Enhancing Learning*, 157–176. https://doi.org/10.4018/979-8-3693-6397-3.ch007

Laohapornchaiphan, J., & Chenprakhon, P. (2024). A Review of Research on Learning Activities Addressing the Submicroscopic Level in Chemistry. *Journal of Chemical Education*, 101(11), 4552-4565. https://doi.org/10.1021/acs.jchemed.4c00156

Morales-Doyle, D., Childress Price, T., & Chappell, M. J. (2019). Chemicals are contaminants too: Teaching appreciation and critique of science in the era of Next Generation Science Standards (NGSS). *Science Education*, 103(6), 1347-1366. https://doi.org/10.1002/sce.21546

Muniroh, N., Rusilowati, A., & Isnaeni, W. (2022). Instrument Development of Science Literacy Assessment with Socio-Sciences Contains in Natural Science Learning for Elementary School. *Journal of Research and Educational Research Evaluation*, 11(1), 15-22. https://doi.org/10.15294/jere.v11i1.55421

OECD (2021), 21st-Century Readers: Developing Literacy Skills in a Digital World, PISA, OECD Publishing, Paris, https://doi.org/10.1787/a83d84cb-en

Pebrianti, P., & Suryani, O. (2024). Development of the Acid-Base Module Based on Problem Based Learning with Ethnochemisty to Improve Students Science Literacy Ability. *Jurnal Penelitian Pendidikan IPA*, 10(8), 4634-4640. https://doi.org/10.29303/jppipa.v10i8.8582

Permanasari, A. (2016, October). STEM education: Inovasi dalam pembelajaran sains. In Seminar Nasional Pendidikan Sains VI 2016. Sebelas Maret University.

Putri, V. L., & Dj, L. (2022). Description of students learning difficulties in the basic law of chemistry. Jurnal Pijar Mipa, 17(5), 597-603. https://doi.org/10.29303/jpm.v17i5.3855

Rahmawati, Y., Ridwan, A., Rahman, A., & Kurniadewi, F. (2019). Chemistry students' identity empowerment through etnochemistry in culturally responsive transformative teaching (CRTT). *Journal of Physics: Conference Series*, 1156, 012032. https://doi.org/10.1088/1742-6596/1156/1/012032

Rosyidi, D. (2020). Teknik dan instrumen asesmen ranah kognitif. *Tasyri': Jurnal Tarbiyah-Syari'ah Islamiyah*, 27(1), 1-13. https://doi.org/10.52166/tasyri.v27i1.79

Rubini, B., Ardianto, D., Pursitasari, I. D., & Hidayat, A. (2018). Science Teachersâ€™ Understanding on Science Literacy and Integrated Science Learning: Lesson from Teachers Training. *Jurnal Pendidikan IPA Indonesia*, 7(3), 259-265. https://doi.org/10.15294/jpii.v7i3.11443

Rusilowati, A., Nugroho, S. E., Susilowati, E. S. M., Mustika, T., Harfiyani, N., & Prabowo, H. T. (2018). The development of scientific literacy assessment to measure student's scientific literacy skills in energy theme. *Journal of Physics: Conference Series*, 983, 012046. https://doi.org/10.1088/1742-6596/983/1/012046

Shwartz, Y., Ben-Zvi, R., & Hofstein, A. (2006). The use of scientific literacy taxonomy for assessing the development of chemical literacy among high-school students. *Chemistry Education Research and Practice*, 7(4), 203-225. https://doi.org/10.1039/B6RP90011A

Ulfah, A. A., Kartono, K., & Susilaningsih, E. (2020). Validity of content and reliability of inter-rater instruments assessing ability of problem solving. *Journal of Research and Educational Research Evaluation*, 9(1), 1-7. https://doi.org/10.15294/jere.v9i1.40423

UNESCO. (2021). Implementation of the UNESCO Recommendation on Open Science. https://www.unesco.org/en/open-science/implementation

Wardani, L. K., Mulyani, B., Ariani, S. R. D., Yamtinah, S., Masykuri, M., Ulfa, M., & Shidiq, A. S. (2023). The effect of an ethnochemistry-based culturally responsive teaching approach to improve cognitive learning outcomes on green chemistry material in high school. *Jurnal Penelitian Pendidikan IPA*, 9(12), 11029-11037. https://doi.org/10.29303/jppipa.v9i12.5532

Zou, Y., Xue, X., Jin, L., Huang, X., & Li, Y. (2024). Assessment of conceptual understanding in student learning of evaporation. *Physical Review Physics Education Research*, 20(2), 020107. https://doi.org/10.1103/PhysRevPhysEducRes.20.020107