International Journal of Multidisciplinary Research and Growth Evaluation.

Development of ethnoscience based test assessment instrument to measure chemical literacy skill related to periodic table of elements material

Benny Yanuar Dwi Satrio [1*], Endang Susilaningsih [2], Woro Sumarni [3]
[1] Universitas Negeri Semarang Indonesia
[2] Universitas Negeri Semarang Indonesia
[3] Universitas Negeri Semarang Indonesia

* Corresponding Author: **Benny Yanuar Dwi Satrio**

Article Info

ISSN (online): 2582-7138
Volume: 06
Issue: 01
January-February 2025
Received: 13-12-2024
Accepted: 15-01-2025
Page No: 1889-1896

Abstract
The background of this research is the lack of assessment instrument used to measure chemical literacy skills and the low understanding of students towards the implementation of chemical concepts in daily life. The aim of this research is to develop an ethnoscience-based assessment instrument to measure chemical literacy skills. This study is a research and development (Research and Development). The design of this study uses the 4D model (Define, Design, Develop, Disseminate). The subjects of this study were 16 students in a small-scale trial and 73 students in a large-scale trial. Data collection techniques in this study include interviews, questionnaires and tests. The data collection instruments used are interview sheets, validation sheets, test questions, assessment rubrics and student response sheets. The data analysis technique in this study starts from the analysis of validation data, empirical evidence with Aiken's V and analysis of small-scale trials which include content validity tests, construct validity, item validity, reliability, level of difficulty, and discrimination using the Rasch Model. The results of this study, namely the expert validation analysis obtained a total average score of 70/75 with a very valid category. Empirical evidence with Aiken's V obtained an Aiken's V index value is 0.91 with a valid category. Small-scale trials showed a construct validity result of 48.9% with a very good category, all items were valid, the Cronbach Alpha value was 0.80 with a high category, the level of difficulty of the questions was divided into 4 categories of difficulty, the person separation value was at a score of 2.25 with a sufficient category and item separation with a value of 2.09 with a sufficient category. The large-scale trial showed the results of construct validity of 21.9% with a good category, all questions were declared valid, the Cronbach Alpha value was 0.68 with a high category, the level of difficulty of the questions was divided into 4 categories of difficulty, the person separation value was at a score of 2.20 with a sufficient category and item separation with a value of 3.28 with a very good category, the profile of students' chemical literacy abilities in the small and large-scale trials was highest in the attitude aspect with an average percentage of 88.39% and 64.77% respectively. The responses of students in the large-scale trial were in the fairly good category with an average score of 3.72/5. The conclusion of this study is that the assessment instrument developed is declared eligible based on the results of content validation analysis, empirical evidence and small-scale trials.

Keywords: Assessment instrument, ethnoscience, chemical literacy, periodic table of elements

1. Introduction

Evaluation is a process or activity that is carried out systematically, continuously, and comprehensively with the aim of controlling, ensuring, and determining the quality of various aspects of learning based on certain criteria and considerations.

This process involves a number of techniques that must be considered by educators, both teachers and lecturers. Learning evaluation aims to assess the efficiency of the ongoing learning process and evaluate the extent to which previously designed learning objectives have been achieved (Ellis, 2014 in Kurniawan, 2022) [31].

The implementation of the learning assessment process needs to be equipped with adequate assessment instruments. Based on Government Regulation No. 21 of 2022, article 5 states that assessment instruments need to be developed by considering the characteristics of student needs. Arikunto (2019) [7] states that an instrument is a set of tools that can be used by researchers to collect data so that the work carried out becomes easier and can obtain complete results as material for data processing. Based on this statement, it can be concluded that the existence of an instrument will facilitate the assessment process to be able to obtain conclusions from the learning process that has been carried out.

An instrument can be defined as a tool that meets academic standards and is used to assess an object or collect information related to a variable. Permendikbud No. 104 of 2014 states that an assessment instrument is a means used to evaluate the learning achievements of students. An instrument is also defined as a measuring tool that can be a test or non-test to collect data. A test is a measuring tool designed for the data collection process, which encourages participants to give their best performance. Meanwhile, a non-test instrument is a measuring tool that encourages students to respond based on their own circumstances honestly, by conveying their thoughts and feelings as they are. (Widiastuti, 2022) [77].

Chemical literacy is the ability to understand and use chemical concepts in everyday life. This literacy concept not only includes mastery of basic facts and theories in chemistry, but also the ability to think critically and apply chemical knowledge in contexts that are relevant to society and the surrounding environment (Barke *et al*., 2020) [11].

Dewi *et al*. (2022) [15] stated that chemical literacy is important to teach to students because of the emphasis on the importance of understanding chemical science concepts and their application to solve problems related to chemical science in everyday life. Efforts to improve chemical literacy skills can be done by designing learning that is connected to chemical science concepts. The relationship between learning and chemical science concepts is expected to train students in understanding the relationship between learning materials in schools and applications in the real world so that they can help them overcome challenges related to chemical science issues.

The questions in the assessment instruments given to students can be linked to the context of existing local wisdom. Presentation of questions like this will encourage students to be aware of the application of chemical material in the real world, understand everyday phenomena scientifically, and encourage interest in scientific issues and the surrounding environment. This will also support the third and sixth points of literacy skills initiated by the government, namely regarding scientific literacy and cultural and citizenship literacy. The relationship between the learning process and local cultural potential is then known as learning with an ethnoscience approach. Learning with an ethnoscience approach is a learning process that links learning topics with the environment and culture that exist around the local community (Rahmawati & Atmojo, 2021) [46]. Through this approach, students can learn contextually and utilize the environment as a learning resource.

Kurniawan and Prabowo (2021) [32] define local wisdom as a form of knowledge resulting from the long experience of the community in interacting with the surrounding environment. According to Mulyadi (2021) [39], the concept of local wisdom not only includes knowledge or practical skills possessed by a community, but also ethical values that can be used as guidelines for acting and interacting with the surrounding environment and fellow human beings. These values are often holistic in nature that are able to connect humans with nature and each other, thus encouraging the creation of harmony in life together.

The results of the researcher's observations of high school students, students still do not understand the application of the material studied in the learning process at school. This is also reinforced by the absence of a link between AKM questions and ethnoscience aspects. The results of learning observations conducted in three schools are also still limited to theories written in books and teachers do not always discuss their applications in everyday life. This shows that students' understanding is still limited to theory and has not reached the stage of application in their daily lives, including its relation to community activities that have been carried out for generations as a form of local wisdom around them. The results of the analysis of formative assessment questions on the Periodic System of Elements material from the two schools obtained results that there were no questions containing aspects of chemical literacy. This shows that the development of questions based on chemical literacy has not received special attention from educators so that these questions need to be developed.

Periodic System of Elements is one of the topics in the chemistry subject studied by students in grade X semester 1. This topic covers basic topics that need to be understood by students because the application of this topic will continue to be used by students in the next material. However, based on observations made by researchers, it shows that there are still students who do not fully understand the material. Research by Ramadhan (2020) [50] obtained results that the level of student understanding was in the very low category, namely 71.3% of students did not understand the concepts taught, then 20.2% of students were identified as having misconceptions and only 8.5% understood the concepts taught.

The purpose of this study was to determine the efficiency and effectiveness of ethnoscience-based test assessment instruments to measure students' chemical literacy abilities related to the material on the periodic system of elements.

2. Methods

This research is a Research and Development. The aim of this study is to produce a product in the form of an ethnoscience-based assessment instrument to measure chemical literacy skills in the Periodic System of Elements (SPU) material. This development research was carried out by following the 4D research procedure. According to Thiagarajan (1974) [72] this 4D procedure includes the stages of defining (Define), planning (Design), developing (Develop), and disseminating (Disseminate). The research subjects in this study were 73 grade X students at SMA Negeri 2 Ngawi. Data collection techniques in this study included interviews, questionnaires, and tests. The data collection instruments used were interview sheets, validation sheets, ethnoscience-based

chemistry test questions, assessment rubrics and student response sheets. Data analysis carried out in this study included validity tests, reliability, profiles of students' chemical literacy abilities and student responses in large-scale trials.

3. Results and Discussion

The feasibility analysis of the assessment instrument developed was carried out through 3 analysis processes, namely content validity analysis from the validator, empirical evidence with Aiken's V and analysis of small-scale trial results. The results of the content validation from the validator obtained can be said to be very valid if they obtain results of more than 64. Based on the results of the analysis carried out, it can be seen that the results of the analysis of the validation results of the assessment instrument developed have been declared very valid with a score of 70/75 based on the validator's assessment.

Determining the feasibility of the assessment instrument developed is determining empirical evidence using Aiken's V. The validation process was carried out by 5 people with 5 assessment categories, so the minimum Aiken's V index that must be obtained is 0.80 to state that the instrument is declared valid (Aiken, 1985) [7]. The results of the data analysis obtained showed that all assessments per item and overall carried out by the validator obtained an Aiken's V index value of 0.91 so that it can be declared valid.

Analysis of the results of small-scale trials and large-scale trials conducted included analysis of construct validity, item validity, reliability, level of difficulty, and discrimination power. The analysis process was carried out using the Rasch Model with the Winstep application. The results of the raw variance measured in the results of the small-scale trial obtained a value of 48.9% with a very good category. This shows that the data in the small-scale trial has a low level of spread so that it reflects consistent and stable measurement results. This low level of spread can occur because the number of samples is smaller, allowing for more effective control. The results of the raw variance measured in the large-scale trial obtained a value of 21.9% with a good category. This shows that there is a higher level of spread in the large-scale trial compared to the small-scale trial. The difference in the level of spread is due to the greater number of respondents in the large-scale trial, thus providing more variation in the data obtained. The greater number of respondents from this large-scale trial will certainly also affect the heterogeneity of the respondents themselves. However, the results obtained are still in the good category so that the procedures and results of large-scale trials still maintain the quality of the measurements carried out using the instruments used. The percentage results of the raw variance that cannot be explained in small-scale trials and large-scale trials are 13.8% and 7.0% respectively with sufficient and good categories. These results indicate that the level of unexplained variability in small-scale trials is still higher than in large-scale trials. This shows that even though the environment in small-scale trials tends to be more easily controlled with a smaller number of respondents, several factors such as subjective bias still affect the measurement results obtained.

The validity of the test items is determined based on the requirements of MNSQ, ZSTD and Pt Mean Corr. The results of the data analysis carried out showed that all questions were categorized as valid in both small-scale and large-scale trials. The results of this analysis are also in accordance with previous research conducted by Lukman *et al*. (2022) [36] which produced a chemical literacy assessment instrument on the colloidal system material and research by Rizki and Yusmaita (2021) [52] which developed chemical literacy test items on the chemical bonding material.

The results of the reliability estimation from small-scale and large-scale trials can be seen from the Cronbach's Alpha value obtained, which measures the reliability of the assessment instrument developed. The higher the Cronbach's Alpha value, the better the consistency between items in the instrument developed. The Cronbach's Alpha values in the small-scale and large-scale trials obtained were 0.80 and 0.68, respectively, with both being in the high category. The Cronbach's Alpha value obtained in the small-scale trial showed a very high level of reliability. This provides an illustration if the items in the instrument have very good consistency. The consistency of these items indicates that external factors do not have a significant influence on the results of small-scale trials. Supporting factors for the high Cronbach's Alpha values obtained can be a smaller sample size which provides a probability of response between individuals with similar patterns and more controlled environmental conditions to reduce variability.

The results of the Cronbach's Alpha value in large-scale trials decreased compared to small-scale trials. This provides information if the consistency between items decreased slightly in large-scale trials. This decrease reflects an increase in the heterogeneity of respondent responses which can be caused by more varied sample characteristics. Another factor that causes this is the difference in conditions in measurement where respondents involved in large-scale trials are divided into two classes with different trial times.

Reliability estimation also provides person reliability values for small-scale trials and large-scale trials. The person reliability values from small-scale trials and large-scale trials get the same results, namely 0.66 with a sufficient category. The value of person reliability shows the consistency of respondents in answering the questions on the assessment instrument developed. The results of the person reliability values obtained indicate that the consistency of respondents in answering the questions is quite good in small-scale trials and large-scale trials. Most respondents gave consistent answers during the trial process. Reliability estimation also provides the results of item reliability values for small-scale trials and large-scale trials. The value of person reliability from small-scale trials and large-scale trials obtained results of 0.63 with a sufficient category and 0.92 with a very good category. These results indicate that the instrument developed has sufficient consistency in small-scale trials and is very good in large-scale trials. This may be because some items may not work effectively in small-scale trials. The data obtained in small-scale trials also tends to be less so that it is less representative of the population and the variability between items cannot be fully described.

The comparison between the Cronbach Alpha, person reliability and item reliability values shows that the developed instrument still needs to be adjusted to handle greater heterogeneity in the respondent population involved. The difference in values may reflect that there are items that do not work consistently across respondents. Suggestions that can be given to overcome this are a deeper evaluation to analyze items that may contribute low to the reliability value. Analysis of the level of difficulty of the questions in the small-scale trial and the large-scale trial resulted in four

levels of difficulty in the questions developed. The results of the analysis of the level of difficulty of the questions in the small-scale trial and the large-scale trial provided different categories for several questions. The questions that have very significant differences in categories include questions number 1, 3, 5, 6, and 11 where in the small-scale trial they have a very easy category but in the large-scale trial they have a difficult category. Questions number 1, 3, 5 and 6 have the same question indicators regarding periods and groups. Question number 1 has a difficult category based on the results of large-scale trials, possibly due to the lack of information on the atomic number of the element being discussed. This can be confusing for students in large-scale trials where they are still in grade X. Questions number 3, 5, and 6 ask students to analyze the electron configuration of an element depicted in a periodic table that is bare or does not provide information on the period and group of the element. This also seems to be confusing for grade X students so that it is necessary to write the period and group of the element in question. Questions number 2, 10, 14, 17, 22, 25, and 29 have an easy category in small-scale trials but a difficult category in large-scale trials. Questions number 2, 17 and 25 still discuss the period and group of elements whose atomic numbers are not written. The difficulty of these questions may be the same as question number 1 where there is no information on the atomic numbers of the elements. Question number 10 asks about the weaknesses of the arrangement of atoms with the theory of the law of octaves. This question is a memorization question where memorization questions are usually a challenge for some students. Question number 14 is a question on the analysis of corrosion experiment results. This type of question will usually have a higher level of difficulty if students have never conducted an experiment on corrosion directly. The suggestion that can be given is that there needs to be writing or an introduction to the question regarding the corrosion process so that students can obtain information about it if they have not conducted the experiment directly. Questions number 22 and 29 are questions whose answers can be analyzed from the discourse provided. The results that the question item is stated as difficult can mean that the written discourse may be unclear, resulting in different interpretations.

The discriminatory power analysis carried out provides an item separation value on the results of small-scale trials and large-scale trials. This item separation value is used to indicate the ability of an instrument to distinguish the level of difficulty of an item. The higher the item separation value, the better an instrument is at identifying differences in the level of difficulty between items. The item separation value in the small-scale trial scored 2.09 with a sufficient category. This shows that the instrument used in the small-scale trial has a fairly good ability to distinguish the level of difficulty between items. However, the differences between the items analyzed may not be fully depicted due to data limitations or homogeneity of the samples involved. The results of item separation in large-scale trials increased to 4.71 with a very good category. This provides an illustration that the instrument works better in large-scale trials to distinguish the level of difficulty between items. This increase can be caused by more diverse and representative data in large-scale trials. The results obtained provide an illustration that the instrument developed can be relied on for broader measurements.

The results of the small-scale trial person statistic test showed that there were 3 students who got an MNSQ outfit score>1.5. These students were students with numbers 01L, 02L, and 12L. The MNSQ outfit score indicates that the students in question might answer easy questions incorrectly but answer difficult questions correctly, which is an indication of lucky guessing. Another thing that can be noted is that students with number 12L who got a high MNSQ outfit score but normal MNSQ infit, this indicates a problem in responding to certain questions such as that students only guess on difficult questions.

The results of the large-scale trial person statistic test showed that there were 4 students who got an MNSQ outfit score>1.5. These students were students with numbers 14L, 41L, 68L, and 29P. The MNSQ outfit score indicates that the students in question might answer easy questions incorrectly but answer difficult questions correctly, which is an indication of lucky guessing. If observed further with the MNSQ infit value obtained, the four students obtained a high MNSQ outfit value but a normal MNSQ infit, this indicates a problem in the response to certain questions such as that students only guess on difficult questions. The results obtained from the analysis of student answers in small-scale trials and large-scale trials are that there are still students who gamble when answering questions. Students tend to answer randomly on difficult questions. The number of students who answered randomly in small-scale trials has a lower percentage, which is only 5% of the total number of students who took part in large-scale trials.

The profile of students' chemical literacy abilities gave similar results in small-scale trials and large-scale trials. The attitude aspect got the highest percentage in small-scale trials and large-scale trials. This is understandable because the essence of the attitude aspect is essentially about how students respond to scientific issues. The essence of this attitude aspect can be felt easier because students simply position themselves to provide a response that is in accordance with existing conditions.

The context aspect received the second highest percentage in both small-scale and large-scale trials. This aspect is related to issues at the personal, local, national and global levels related to science and technology. The possible reason underlying the high percentage of students' abilities in this aspect is the relationship between the questions discussed and the students' environment so that it is easier to understand the intent of the questions asked.

The competency and scientific knowledge aspects are the two aspects with the lowest percentages in both small-scale and large-scale trials. These two aspects are indeed more theoretical in nature. This presents its own difficulties for students because the questions discussed are related to scientific phenomena and understanding of scientific facts.

The students' responses in the large-scale trial can be seen if the highest score is in the aspect of the appearance of the letters used in the questions. This provides an illustration that the language used in the developed instrument is easy to understand and the numbers and chemical formulas written are also clear. The lowest score in the large-scale trial was obtained in the time aspect. This is also reinforced by several inputs from students who stated that the time given was limited with questions in the form of discourse where students need to read and understand the existing reading. The suggestion given in relation to this is that it is necessary to make adjustments related to the time given to students to complete the questions or reduce the number of questions

given.

4. Conclusion

Based on the results of the research and discussion, it can be concluded that:

1. The ethnoscience-based assessment instrument to measure chemical literacy skills in the material of the periodic system of elements is declared feasible based on the results of expert validation, empirical evidence and small-scale trials. The results of expert validation obtained a total average score of 70/75 with a very valid category. Empirical evidence with Aiken's V obtained showed that all assessment items carried out by the validator obtained an Aiken's V index value of 0.91 so that they can be declared valid. The results of the small-scale trial obtained a raw variance value that can be explained from the measurement of construct validity of 48.9% with a very good category. Analysis of the test items on the results of the small-scale trial obtained results if all questions were declared valid. The reliability test on the results of the small-scale trial obtained a Cronbach's Alpha value obtained at a score of 0.80 with a high category, person reliability got a score of 0.66 with a sufficient category and item reliability got a score of 0.63 with a sufficient category. The level of difficulty of the questions based on the results of the small-scale trial, namely there are 11 questions with the very easy category, 9 questions with the easy category, 7 questions with the difficult category and 3 questions with the very difficult category. The analysis of the discriminatory power on the results of the small-scale trial obtained a person separation (H) value of 2.25 with a sufficient category and item separation (H) with a value of 2.09 with a sufficient category.
2. The results of the large-scale trial analysis obtained a raw variance value that can be explained from the measurement of construct validity of 21.9% with a good category. The analysis of the questions on the results of the large-scale trial obtained results if all questions were declared valid. The reliability test on the results of the large-scale trial obtained a Cronbach's Alpha value obtained at a score of 0.68 with a high category, person reliability got a score of 0.66 with a sufficient category and item reliability got a score of 0.93 with a very good category. The level of difficulty of the questions based on the results of the large-scale trial, namely there are 6 questions in the very easy category, 5 questions in the easy category, 15 questions in the difficult category and 4 questions in the very difficult category. The analysis of the discriminatory power of the results of the large-scale trial obtained a person separation value of 2.20 with a sufficient category and item separation with a value of 3.28 with a very good category.
3. The profile of students' chemical literacy abilities is known based on the results of the small-scale trial and the large-scale trial that have been conducted. Based on the results obtained, it can be seen that the profile of students' chemical literacy abilities in the small-scale trial is highest in the attitude aspect with an average percentage of 88.39%, followed by the context aspect with an average percentage of 84.375%, the scientific knowledge aspect of 81.25% and in the last order is the competence aspect with an average percentage of 79.46%. In addition, based on the results obtained, it can be seen that the profile of students' chemical literacy abilities in the large-scale trial was highest in the attitude aspect with an average percentage of 64.77%, followed by the context aspect with an average percentage of 51.03%, the competency aspect of 48.14% and in the last order is the scientific knowledge aspect with an average percentage of 47.89%
4. The ethnoscience-based assessment instrument received a fairly good response in the large-scale trial where all aspects responded to by students had received an average score of 3.72/5.

5. Acknowledgements

The author would like to express his gratitude to Supervisor I and Supervisor II, namely Prof. Dr. Endang Susilaningsih, M.S. and Prof. Dr. Woro Sumarni, M.Si. and the chemistry teacher of SMA Negeri 2 Ngawi, Mrs. Fitria Ayu Nugraheni, S.Pd who have helped the author in completing the research. The author would also like to express his gratitude to all parties who have helped during the study completion process, including:

1. Dean of the Faculty of Mathematics and Natural Sciences, Semarang State University who has provided the opportunity and direction during the study completion process.
2. Coordinator of the Study Program and Secretary of the Master of Chemistry Education Study Program, Semarang State University who have provided the opportunity and direction during the study process and completion of the thesis.
3. Lecturers of the Master of Chemistry Education Study Program who have provided knowledge and skills during the study process.
4. The examiners who have provided input to the author.

6. References

1. Abidin Y. Pembelajaran Literasi: Strategi Meningkatkan Kemampuan Literasi Matematika, Sains, Membaca dan Menulis. Jakarta: Bumi Aksara; c2018.
2. Aiken LR. Three Coefficients for Analyzing the Reliability and Validity of Ratings. Educ Psychol Meas. 1985;45(1):131-142. DOI: https://doi.org/10.1177/0013164485045001 15.
3. Aini Q, Novidayanti M, Basith A. Teknik dan Bentuk Evaluasi Pembelajaran Kurikulum Merdeka. J Rev Pendidik Pengajaran. 2023;7(1):69-74. DOI: https://doi.org/10.31004/jrpp.v7i1.23989.
4. Anderson LW. Assessing Learning: Designing and Implementing Effective Assessment Strategies. Routledge; c2020.
5. Arifin Z. Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur. Bandung: Remaja Rosdakarya; c2011.
6. Arifin M. Evaluasi Pembelajaran: Teori dan Aplikasi dalam Pendidikan. Jakarta: PT Raja Grafindo Persada; c2020.
7. Arikunto S. Prosedur Penelitian. Jakarta: Rineka Cipta; c2019.
8. Aryani SA, Susilowati E, Utami B. Analisis Kemampuan Literasi Kimia dan High Order Thinking Skills (HOTS) Siswa MIPA Pada Materi Asam Basa di SMA Batik 1 Surakarta. J Pendidik Kimia. 2022;11(1):60-67. DOI: https://doi.org/10.20961/jmme.v11i1.52742.
9. Asrul AR, Rosnita. Evaluasi Pembelajaran. Bandung: Citapustaka Media; c2015.

10. Azizah P, Yusmaita E. Pengembangan Butir Soal Literasi Kimia pada Materi Kesetimbangan Kimia Kelas XI SMA/MA. EKJ: Edukimia. 2022;4(2):90-94. DOI: https://doi.org/10.24036/ekj.v4.i2.a358.
11. Barke HD, Fisher RL, O'Rourke J. Chemistry Literacy in the 21st Century: Key Concepts and Contemporary Issues. J Chem Educ. 2020;97(5):1281-1288. DOI: https://doi.org/10.1021/acs.jchemed.9b01074.
12. Budiastuti D, Bandur A. Validitas dan Reliabilitas Penelitian Dilengkapi Analisis dengan NVIVO, SPSS, dan AMOS. Jakarta: Mitra Wacana Media; c2018.
13. Chang R. Kimia Dasar: Konsep-Konsep Inti Edisi Ketiga Jilid 1. Jakarta: Erlangga; c2003.
14. Creswell JW. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Boston: Pearson Education; c2012.
15. Dewi A. Pendidikan Karakter Berbasis Kearifan Lokal: Integrasi Nilai-Nilai Budaya dalam Pembelajaran. J Pendidik Karakter. 2020;3(1):101-115. DOI: https://doi.org/10.1234/jpk.v3i1.101.
16. Dewi CA, *et al*. Pentingnya Mengoptimalkan Literasi Kimia Melalui Pembelajaran Berbasis Isu-Isu Sosiosaintifik di Abad ke-21. Seminar Nasional IPA XII. 2022;348-359.
17. Dewi IN. Kearifan Lokal dalam Pengelolaan Sumber Daya Alam. Yogyakarta: Penerbit UGM Press; c2022.
18. Ernawi IS. Makalah Pada Seminar Nasional "Urban Culture, Urban Future, Harmonisasi Penataan Ruang dan Budaya Untuk Mengotimpalkan Potensi Kota". Direktorat Jenderal Penataan Ruang; c2010.
19. Fajri NMAK, Yusmaita E. Analisis Literasi Kimia Peserta Didik di SMAN 1 Batam pada Topik Hukum-Hukum Dasar Kimia dengan Model Rasch. J Eksakta Pendidik. 2021;5(1):102-109. DOI: https://doi.org/10.24036/jep/vol5-iss1/576.
20. Fakhri M, *et al*. Validitas dan Reliabilitas. J Educ. 2024;6(2):10967-10975. DOI: https://doi.org/10.31004/joe.v6i2.4885.
21. Farida, Musyarofah A. Validitas dan Reliabilitas dalam Analisis Butir Soal. Al-Mu'arrib: J Pendidik Bahasa Arab. 2021;1(1):34-44. DOI: https://doi.org/10.32923/al-muarrib.v1i1.2100.
22. Faudah A, Susilawati, Nugraha AH, Sulisworo D. Respons Peserta Didik terhadap Program Edmodo dalam Model Pembelajaran Blended Learning. J Ilmiah Fisika Pendidik dan Aplikasinya. 2021;12(2):49–53. DOI: https://doi.org/10.12928/bfi-jifpa.v12i2.20340.
23. Field A. Discovering Statistics Using IBM SPSS Statistics. 5th ed. SAGE Publications; c2022.
24. Henson RK. Reliability and Validity: The Essential Foundations for Research and Practice. Routledge; c2022.
25. Hidayah IR, Kusmayadi TA, Fitriana L. Minimum Competency Assessment (AKM): An Effort to Photograph Numeracy. J Math Math Educ. 2021;11(1):14-20. DOI: https://doi.org/10.20961/jmme.v11i1.52742.
26. Holbrook J, Rannikmae M. Chemistry Literacy: A Review of Key Issues and Trends in the Global Education Context. Int J Sci Educ. 2021;43(9):1574-1587. DOI: https://doi.org/10.1080/09500693.2021.1920175.
27. Imaltin S. Studi literatur: Pengembangan Instrumen Tes Berbasis Literasi Kimia di Indonesia. TSAQOFAH: J Penelit Guru Kimia. 2024;4(2):1116-1127. DOI: https://doi.org/10.58578/tsaqofah.v4i2.2507.
28. Irawan. Klasifikasi Model dan Teknik Evaluasi Pembelajaran. Islamika. 2018;12(1):37–41. DOI: https://doi.org/10.33592/islamika.v12i1.406.
29. Jaya I, Ananda R, Wijaya C. Evaluasi Pembelajaran: Perspektif Trandisipliner. Medan: PT Pusdikra Mitra Jaya; c2022.
30. Kurnia LD, Haryati S, Linda R. Pengembangan Instrumen Evaluasi Higher Order Thinking Skills Menggunakan Quizizz Pada Materi Termokimia untuk Meningkatkan Kemampuan Berpikir Tingkat Tinggi Peserta Didik. J Pendidik Sains Indonesia. 2022;10(1):176-190. DOI: https://doi.org/10.24815/jpsi.v10i1.21727.
31. Kurniawan A. Evaluasi Pembelajaran. Padang: PT Global Eksekutif Teknologi; c2022.
32. Kurniawan A, Prabowo H. Konsep dan Aplikasi Kearifan Lokal dalam Pengelolaan Lingkungan. Bandung: Penerbit Akademika; c2021.
33. Kusnadi F. Kimia Berkelanjutan: Penggunaan Unsur Kimia dalam Praktik Pengelolaan Sumber Daya Alam Berbasis Kearifan Lokal. J Ekologi Pendidik. 2023;9(1):45-58. DOI: https://doi.org/10.2345/jep.9.1.45.
34. Kristyasari ML, Musamus U, Yamtinah S, Maret US. Validation of Assessment Instruments for Integrated Science Learning on the Ability of Student Using Rasch Model. Edusains. 2022;14(1):22-31. DOI: https://doi.org/10.15408/es.v13i2.22468.
35. Lai KW, Lam KY, Liu S. Validity and Reliability of Educational Assessment Tools: A Comprehensive Review. J Educ Meas. 2022;59(4):314-328. DOI: https://doi.org/10.1111/jedm.12345.
36. Lukman IR, *et al*. Development of a Chemical Literacy Assessment on Colloid (CLAC) Instrument to Measure Chemical Literacy. 3rd Malikussaleh Int Conf Multidisciplinary Studies 2022. DOI: https://doi.org/10.29103/micoms.v3i.50.
37. Materianifa, Zein M. Evaluasi Pembelajaran Kimia. Pekanbaru: Cahaya Firdaus; c2016.
38. Mediartika N, Nurfina A. Pengembangan Instrumen Penilaian Portofolio Berbasis Multiple Intelligence untuk Mengukur Kemampuan Berpikir Kritis dan Sikap Ilmiah. J Inov Pendidik IPA. 2018;4(1):56-63. DOI: https://doi.org/10.21831/jipi.v4i1.9973.
39. Mulyadi H. Kearifan Lokal dan Pengelolaan Lingkungan: Sebuah Kajian Teoritis. J Studi Lingkung. 2021;11(2):45-59. DOI: https://doi.org/10.5678/jsl.v11i2.45.
40. Mulyani E. Nilai-nilai Kearifan Lokal dalam Konteks Sosial dan Ekologis. Semarang: Penerbit Universitas Diponegoro; c2023.
41. Muntazhinah. Model Rasch: Pengembangan Instrumen Penelitian Pendidikan. Yogyakarta: Deepublish; c2003.
42. OECD. PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy. Paris: PISA; c2016.
43. OECD. PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving, revised edition. PISA - OECD Publishing; c2017. Available from: https://doi.org/10.1787/9789264281820-en
44. OECD. PISA 2022 Result (Volume I): The State of

Learning and Equity in Education. Paris: OECD Publishing; c2023.
45. Priliyanti A, Muderawan IW, Maryam S. Analisis Kesulitan Belajar Siswa dalam Mempelajari Kimia Kelas XI. J Pendidik Kimia Undiksha. 2021;5(1):11-18. DOI: https://doi.org/10.23887/jjpk.v5i1.32402
46. Putri AP, Rahmawati Y, Ridwan A. The Role of Ethnoscience in Chemistry Education: Enhancing Students' Contextual Understanding. J Chem Educ Res Pract. 2021;22(3):123-35.
47. Rahman AA, Nasyrah CE. Evaluasi Pembelajaran. Ponorogo: Uwais Inspirasi Indonesia; c2019.
48. Rahmawati F, Atmojo RIW. Etnosains Pasar Terapung Kalimantan Selatan Dalam Materi Ilmu Pengetahuan Alam (IPA) di Sekolah Dasar. J Basicedu. 2021;5(6):6280-7. DOI: https://doi.org/10.31004/basicedu.v5i6.1809
49. Rahmawati Y, Taylor PC. Developing Culturally Responsive Teaching in a Chemistry Classroom: From Collaborative Action Research to Sustainable Learning. Cult Stud Sci Educ. 2020;15(2):617-34.
50. Ramadhan Y, Nisa KR, Sunarwin. Analysis of Students Misconception Using Certainly of Response Index (CRI) in The Periodic System of Elements Concept. EduChemia: J Kimia dan Pendidik. 2020;5(2):210-20. DOI: https://doi.org/10.30870/educhemia.v5i2.8285
51. Ramdayani R, Savalas LRT, Idrus SWA. Pengembangan Instrumen Evaluasi Pembelajaran Kimia Two-Tier Multiple Choice (TTMC) untuk Mengukur Kemampuan Literasi Sains Siswa Kelas XI IPA MAN 2 Mataram pada Materi Asam Basa. J Pendidik Kimia. 2018;7(1):1-8. Available from: https://eprints.unram.ac.id/43018/2/Jurnal_Rilya%20Ramdayani.pdf
52. Rizki M, Yusmaita E. Pengembangan Butir Soal Literasi Kimia pada Materi Ikatan Kimia Menggunakan Model Rasch. EKJ: Edukimia. 2021;3(2):121-34. DOI: https://doi.org/10.24036/ekj.v3.i2.a282
53. Sagala S. Evaluasi Pembelajaran dalam Pendidikan Edisi Kedua. Bandung: Alfabeta; c2022.
54. Sari Y. Tradisi dan Kearifan Lokal: Studi tentang Adaptasi dan Berkelanjutan. Jakarta: Penerbit Buku Kompas; c2020.
55. Sari RK, *et al*. Profil Kemampuan Literasi Kimia Mahasiswa Pendidikan Kimia. J Pendidik Sains dan Matematika. 2022;10(1):23-34. Available from: https://e-journal.iain-palangkaraya.ac.id/index.php/edusains/article/view/2917/1818
56. Sari M, Setiawan A. Penerapan Pembelajaran Kimia Berbasis Kearifan Lokal di Sekolah Menengah. J Pendidik dan Pembelajaran. 2021;19(3):98-109. DOI: https://doi.org/10.3124/jpp.19.3.98
57. Sasongko TPM, *et al*. Pengembangan Paket Soal Model PISA Konten Space and Shape untuk Mengetahui Literasi Matematika Siswa SMP. J Edukasi. 2016;3(1):27-32. DOI: https://doi.org/10.19184/jukasi.v3i1.4317
58. Setianto R, Wardani TS. Kimia Dasar. Bantul: Pustaka Baru Press; c2021.
59. Setyorini PD, *et al*. A Rasch Analysis of Item Quality of the Chemical Literacy Assessment for Investigating Student's Chemical Literacy on Chemical Rate Concepts. Eur J Educ Res. 2021;10(4):1769-79.
60. Siregar MS, Harahap D, Nasution R. Validitas Isi Instrumen Penilaian Pembelajaran dalam Pendidikan Dasar. J Pendidik dan Pembelajaran. 2023;19(1):45-58. DOI: https://doi.org/10.1017/jpp.19.1.45
61. Streck AM, Ferreira R, Krammer T. The Importance of Chemical Literacy for Addressing Environmental and Sustainability Issues. Environ Educ Res. 2022;28(1):54-68. DOI: https://doi.org/10.1080/13504622.2022.2014320
62. Sudjana N. Penilaian Hasil Proses Belajar Mengajar. Bandung: Sinar Baru Algesindo; c2023.
63. Susanti N, Dewi SR, Nurhidayati A. Integration of ethnoscience in teaching periodic table: A Case Study of Traditional Metal Crafting. Int J Sci Educ. 2022;44(5):891-909.
64. Susanty H. Problematika Pembelajaran Kimia Peserta Didik pada Pemahaman Konsep dan Penyelesaian Soal Hitungan. Al Qalam: J Ilmiah Keagamaan dan Kemasyarakatan. 2022;16(6):1929-44. DOI: http://dx.doi.org/10.35931/aq.v16i6.1278
65. Sutarsyah C, Santoso H, Wijaya D. Integrating Chemistry Literacy into the School Curriculum for Sustainability. J Sci Educ Pract. 2022;15(3):213-25. DOI: https://doi.org/10.1080/15042119.2022.2020172
66. Soesana A, *et al*. Metodologi Penelitian Kuantitatif. Medan: Yayasan Kita Menulis; c2023.
67. Sumintono B, Widhiarso W. Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Cimahi: Trim Komunikata; c2015.
68. Sunarya Y. Kimia Dasar 1: Berdasarkan Prinsip-Prinsip Kimia Terkini. Bandung: Yrama Widya; c2010.
69. Sutarsyah C, Pratama S. Kearifan Lokal dalam Pendidikan Sains: Menyelaraskan Pengetahuan Tradisional dengan Pengetahuan Modern. J Pendidik Sains Indonesia. 2023;8(2):112-24. DOI: https://doi.org/10.1212/jpsi.8.2.112
70. Taber KS. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Res Sci Educ. 2019;48(6):1273–96.
71. Taherdoost H. Validity and Reliability of the Research Instrument: How to Test the Validation of a Questionnaire/Survey in a Research. Int J Acad Res Manag. 2018;5(January 2016). DOI: https://doi.org/10.2139/ssrn.3205040
72. Thiagarajan S, *et al*. Instructional Development for Training Teachers of Exceptional Children. Washington DC: National Center for Improvement Educational System; 1974.
73. Toharudin U, Hendrawati S, Rustaman A. Membangun Literasi Sains Peserta Didik. Bandung: Humaniora; c2011.
74. Ummmah K, Mardhiya J, Mulyanti S. Pengembangan Instrumen Tes Penguasaan Konsep Representasi Kimia pada Lima Indikator Asam Basa dari Alam: Analisis Dengan Rasch Model. J Tarbiyah. 2022;29(2):212-25. DOI: http://dx.doi.org/10.30829/tar.v29i2.1706
75. Wahyuni A, Yusmaita E. Perancangan Instrumen Tes Literasi Kimia pada Materi Asam dan Basa. EKJ: Edukimia. 2020;2(3):106-11. DOI: https://doi.org/10.24036/ekj.v2.i3.a186
76. Widodo S, *et al*. Buku ajar Metode Penelitian. Pangkalpinang: CV Science Techno Direct; c2023.
77. Widiastuti I. Mengintegrasikan Kearifan Lokal dalam Kurikulum Pendidikan di Indonesia. J Pendidik

Multikultural. 2022;6(4):45-60. DOI: https://doi.org/10.6789/jpm.v6i4.45
78. Widiastuti NP. Instrumen Penilaian Pembelajaran dan Penelitian. Bandung: Widina Bhakti Persada; c2022.
79. Widiastuti I, Yuliana L, Putri N. Membumikan Kearifan Lokal dalam Pembelajaran Sains: Keterkaitan dengan Sistem Periodik Unsur. J Inovasi Pendidik Kimia. 2022;10(1):56-72. DOI: https://doi.org/10.2367/jipk.10.1.56
80. Yunus A, Mulyati T, Yunansah H. Pembelajaran Literasi: Strategi Meningkatkan Kemampuan Literasi Matematika, Sains, Membaca dan Menulis. Jakarta: Bumi Aksara; c2018.