JIPK 19 (2) (2025)

Jurnal Inovasi Pendidikan Kimia

https://journal.unnes.ac.id/journals/JIPK

Content Validity Analysis of AKM Reading Literacy Test Instruments on Acid–Base Theory and Equilibrium Using the Aiken's V Method

Dian Pratiwi[1]✉, Sudarmin[1], Sri Yamtinah[2]

[1]Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia.
[2]Chemistry Education, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Surakarta, Indonesia

Article Info

Received: 30-05-2024

Accepted: 03-07-2025

Published: 31-01-2026

Keywords:
AKM reading literacy
Content validity
Aiken's V
Acid–base equilibrium
Chemistry assessment

✉*Corresponding author:*
dianpratiwi20@students.unnes.ac.id

Abstract

The quality of educational assessment plays a crucial role in measuring students' competencies in the 21st century, particularly reading literacy as assessed through the Minimum Competency Assessment (AKM). This study aimed to analyze the content validity of an AKM reading literacy test instrument on acid–base theory and equilibrium using Aiken's V method. The instrument was developed in accordance with the AKM framework, covering various item formats and cognitive levels, and consisted of 30 test items supported by six contextual reading passages. Content validity was evaluated through expert judgment involving nine validators, comprising five university lecturers and four senior high school chemistry teachers. The validators assessed each item based on six criteria: relevance to content, clarity of item formulation, functionality of answer options, alignment with AKM reading literacy components, appropriateness of completion time, and language accuracy. Quantitative data were analyzed using Aiken's V formula, while qualitative feedback was used to refine the test items. The results showed that all 30 items obtained Aiken's V values equal to or exceeding the critical value of 0.71, indicating acceptable content validity. One item demonstrated high validity, while the remaining items exhibited moderate validity. Expert feedback further contributed to improving clarity, contextual accuracy, and alignment with acid–base equilibrium concepts. Overall, the findings confirm that the developed instrument is content-valid and suitable for measuring students' reading literacy in chemistry. Future research should focus on empirical validation to assess reliability and construct validity.

p-ISSN 1979-0503
p-ISSN 2503-1244

INTRODUCTION

Improving the quality of education has become a central focus of sustainable development, as education must continuously adapt to the demands of a rapidly changing world. This effort is essential to prepare students to contribute successfully to the workforce and the global economy in the twenty-first century (Bao & Koenig, 2019). Enhancing educational quality must be accompanied by improvements in assessment practices that are capable of supporting and personalizing learning (Pinto & Leite, 2020). One strategic initiative implemented in Indonesia to address this need is the *Minimum Competency Assessment* (Asesmen Kompetensi Minimum, AKM), which aims to measure students' fundamental competencies in literacy and numeracy.

Reading literacy is a core competency assessed in AKM. Reading literacy extends beyond basic reading skills and encompasses the ability to understand, analyze, evaluate, and reflect on information presented in written texts (OECD, 2019). It requires students to integrate textual information with prior knowledge, critically examine the validity of arguments, and reflect on the meaning conveyed by a text (Liu et al., 2022). In the context of chemistry education, reading literacy is particularly important for understanding complex concepts, such as acid–base theory and equilibrium. The topic of acid–base theory and equilibrium includes subtopics such as acids and bases, salt hydrolysis, and buffer solutions. These topics align with the core principles of content selection used in PISA, as they are contextualized and relevant to everyday life. In addition to conceptual understanding, these topics involve scientific process skills that require higher-order cognitive engagement (Andriani et al., 2019). A comprehensive understanding of these concepts is necessary for students to apply chemical knowledge meaningfully in practical and real-world situations. Consequently, these topics demand strong reading literacy skills, including the abilities to analyze, interpret, and critically evaluate scientific information.

Despite its importance, students' reading literacy skills within specific scientific disciplines have not been adequately measured. Accurate measurement of reading literacy requires assessment instruments that are both valid and reliable. Content validity is one of the most critical aspects in the development of test instruments (Sudaryono et al., 2019). It serves as a fundamental source of evidence in test construction, indicating the extent to which test items adequately represent and align with the intended construct (Delgado-Rico et al., 2015). One effective method for evaluating content validity is Aiken's V method. Aiken's V is a quantitative technique used to assess content validity based on expert judgments (Dewi & Prasetyo, 2016). This method is widely recognized for its simplicity, ease of calculation, and clarity in interpretation, making it suitable for analyzing expert responses in instrument validation studies (Castro Benavides et al., 2022). Through expert evaluation, test instruments can be refined so that only items with high relevance to the measurement objectives are retained. This process not only enhances the content validity of the instrument but also ensures that the assessment accurately and comprehensively measures students' reading literacy skills. Accordingly, this study aims to analyze the content validity of a reading literacy test instrument on acid–base theory and equilibrium using Aiken's V index. The content validity analysis was conducted through expert judgment involving nine evaluators. Each evaluator was provided with a complete validation package consisting of test blueprints, scoring rubrics, test items, scoring guidelines, answer keys, and validation sheets. The evaluators assessed the alignment between item indicators and reading literacy indicators, the appropriateness of conceptual content, the clarity and accuracy of wording, and the formulation of each test item.

METHODS

This study employed a test instrument research and development approach focusing on reading literacy in the topic of acid–base theory and equilibrium. The primary objective of the study was to determine the content validity of the developed reading literacy test instrument. A descriptive quantitative method was applied, using content validity data calculated through Aiken's V formula. The content validity data were obtained from the validation results of the AKM reading literacy test instrument on acid–base theory and equilibrium, which was evaluated by nine experts. The panel of experts consisted of five university lecturers and four senior high school chemistry teachers from Semarang, Indonesia. The test instrument was designed and constructed in accordance with the AKM framework, which includes the distribution of the number of items, the proportional distribution of item formats, and the proportional distribution of cognitive levels. The instrument comprised 30 test items supported by six reading passages that served as stimuli. The distribution of each component is presented in Table 1 and Table 2.

Table 1. Distribution of Item Formats

Item Format	Item Numbers
Simple Multiple Choice	6, 18, 27, 28, 30
Complex Multiple Choice	3, 4, 8, 9, 12, 13, 14, 16, 19, 21, 22, 25, 26
Matching	11
Short Answer	2, 7, 15
Essay	1, 5, 10, 17, 20, 23, 24, 29

Table 2. Distribution of Cognitive Levels of Test Items

Cognitive Level	Item Numbers	Number of Items	Percentage
Locating Information	2, 3, 4, 8, 9, 13, 18, 22, 23	9	30%
Understanding Information	1, 7, 11, 12, 14, 15, 16, 19, 25, 27, 28, 30	12	40%
Evaluating and Reflecting on Information	5, 6, 10, 17, 20, 21, 24, 26, 29	9	30%

Content validity was examined by assigning scores ranging from 1 to 4 to each test item based on the evaluated aspects. The aspects assessed by the experts included: (1) the relevance of the item to the content material; (2) clarity of item formulation to avoid ambiguity; (3) the functionality and plausibility of the answer options; (4) the alignment of the item with the AKM reading literacy components; (5) the appropriateness of the allotted time for item completion; and (6) the correctness and standardization of the language used.

The content validity analysis of the AKM reading literacy instrument in this study employed Aiken's V formula (Aiken, 1985), expressed as follows:

$$V = \frac{\sum s}{n(c-1)}$$

where:

$s = r - l_0$

l_0= the lowest validity rating score (1)

c= the highest validity rating score (4)

r= the score assigned by the expert

n= the number of raters

The quantitative data obtained from the expert ratings were analyzed using Aiken's formula, while the experts' suggestions and comments were analyzed qualitatively to revise and improve the test items. According to Aiken, the minimum acceptable V index value when involving nine experts, six rating categories, and a significance level of 0.05 is 0.71. Therefore, a test item is considered valid if it has an Aiken's V index value of ≥ 0.71.

RESULT AND DISCUSSION

Content validity was established through a comprehensive review of the entire set of test instruments, including the test blueprint, question sheets, scoring rubrics, and answer explanations. This review aimed to ensure that the instrument adequately represented and proportionally reflected the full scope of the content or material to be measured. The content validity assessment of the AKM reading literacy test instrument on acid–base theory and equilibrium was conducted using Aiken's validity index. The results of the content validity analysis based on evaluations from nine experts, calculated using Aiken's formula, are presented in Table 3.

Content validity refers to the degree of alignment between test items and the indicators of AKM reading literacy skills. In this study, six assessment criteria were used to determine content validity, namely: (1) the relevance of each item to the subject matter; (2) clarity of item formulation without ambiguity; (3) the functionality of answer options; (4) the relevance of items to AKM reading literacy components; (5) appropriateness of the time allocation for answering the items; and (6) the use of standardized and proper language.

According to Aiken, with nine raters and six assessment criteria, an item is considered valid if it obtains an Aiken's V index greater than or equal to 0.71. Aiken's V index is a measure of the level of agreement among experts regarding the relevance of test items to the indicators being measured (Cynthia et al., 2023). Based on the results presented in the table, all 30 reading literacy items obtained Aiken's V values higher than the critical V value. Therefore, it can be concluded that all developed reading literacy items are categorized as valid.

The number of experts involved in this study was determined in accordance with the requirements of Aiken's content validity method. One of the key factors influencing the Aiken's V value is the number of validators and the number of assessment criteria used. As the number of validators and criteria increases, the minimum threshold for validity decreases (Kristiyanto et al., 2019). Aiken (1985) classifies content validity into three levels: low validity if the Aiken's V value is below 0.40, moderate validity if the value ranges between 0.40 and 0.80, and high validity if the value exceeds 0.80. Based on the results shown in Table 1, the Aiken's V analysis of the AKM reading literacy instrument indicates that one item falls into the high validity category, while the

remaining 29 items are classified as having moderate validity. The closer the Aiken's V value is to 1, the higher the relevance of the item to the intended indicator (Retnawati, 2016).

Table 3. Results of Aiken's Content Validity Analysis

Item	Aiken's V	Critical V Value	Category
1	0.73	0.71	Valid
2	0.76	0.71	Valid
3	0.80	0.71	Valid
4	0.73	0.71	Valid
5	0.73	0.71	Valid
6	0.80	0.71	Valid
7	0.76	0.71	Valid
8	0.76	0.71	Valid
9	0.73	0.71	Valid
10	0.76	0.71	Valid
11	0.78	0.71	Valid
12	0.80	0.71	Valid
13	0.82	0.71	Valid
14	0.76	0.71	Valid
15	0.73	0.71	Valid
16	0.76	0.71	Valid
17	0.73	0.71	Valid
18	0.73	0.71	Valid
19	0.76	0.71	Valid
20	0.73	0.71	Valid
21	0.73	0.71	Valid
22	0.73	0.71	Valid
23	0.73	0.71	Valid
24	0.76	0.71	Valid
25	0.73	0.71	Valid
26	0.73	0.71	Valid
27	0.76	0.71	Valid
28	0.76	0.71	Valid
29	0.73	0.71	Valid
30	0.80	0.71	Valid

In addition to quantitative analysis, qualitative feedback in the form of suggestions and comments was provided by the nine experts. The main recommendations included: (1) simplifying and clarifying the reading passages without reducing their essential meaning; (2) including sources or references in the passages; (3) ensuring that the passages align with acid–base theory and equilibrium content; (4) preventing potential misconceptions; (5) ensuring that the wording of passages and items complies with Indonesian spelling standards (PUEBI); (6) providing proper numbering and captions for images and graphs; (7) aligning questions and answer options with the given passages; and (8) ensuring that indicators and questions are relevant and consistent with AKM reading literacy components.

Based on the quantitative results, all 30 reading literacy items were declared valid, as their Aiken's V values exceeded the critical threshold. However, the qualitative feedback indicates that revisions are still necessary to improve item quality. These revisions were carried out in accordance with the experts' recommendations. After revising the test items, the instrument development process proceeded to an empirical trial involving senior high school students to examine the empirical validity of the instrument.

CONCLUSION

This study aimed to examine the content validity of an AKM reading literacy test instrument on acid–base theory and equilibrium using Aiken's V method. The results of the content validity analysis involving nine experts

demonstrated that all 30 test items achieved Aiken's V values equal to or greater than the minimum threshold of 0.71, indicating that the instrument meets the criteria for content validity. Specifically, one item was classified as having high validity, while the remaining items were categorized as having moderate validity. These findings indicate that the developed instrument adequately represents the constructs of AKM reading literacy and aligns well with the intended indicators, content scope, and cognitive demands. Furthermore, qualitative feedback from expert validators provided valuable insights for refining item clarity, language use, contextual relevance, and alignment with acid–base theory and equilibrium concepts. The incorporation of these suggestions contributed to improving the overall quality of the instrument. In conclusion, the AKM reading literacy test instrument developed in this study is content-valid and suitable for measuring students' reading literacy skills in the context of acid–base theory and equilibrium. Following the revision process, the instrument is ready to be subjected to empirical testing to further evaluate its psychometric properties, including empirical validity and reliability.

REFERENCES

Aiken, L. (1985). Educational and Psychological Measurement. Pearson.

Andriani, M., Muhali, M., & Dewi, C. A. (2019). Pengembangan Modul Kimia Berbasis Kontekstual Untuk Membangun Pemahaman Konsep Siswa Pada Materi Asam Basa. Hydrogen: Jurnal Kependidikan Kimia, 7(1), 25. https://doi.org/10.33394/hjkk.v7i1.1653

Bao, L., & Koenig, K. (2019). Physics Education Research for 21st Century Learning. Disciplinary and Interdisciplinary Science Education Research, 1(1), 1–12. https://doi.org/10.1186/s43031-019-0007-8

Castro Benavides, L. M., Tamayo Arias, J. A., Burgos, D., & Martens, A. (2022). Measuring digital Transformation in Higher Education Institutions – Content Validity Instrument. Applied Computing and Informatics. https://doi.org/10.1108/ACI-03-2022-0069

Cynthia, C., Arafah, K., & Palloan, P. (2023). Development of Interactive Physics E-Module to Improve Critical Thinking Skills. Jurnal Penelitian Pendidikan IPA, 9(5), 3943–3952. https://doi.org/10.29303/jppipa.v9i5.2302

Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2015). Content Validity Evidences in Test Development: An Applied Perspective. International Journal of Clinical and Health Psychology, 12(3), 449–460.

Dewi, N. D. L., & Prasetyo, Z. K. (2016). Pengembangan Instrumen Penilaian IPA untuk Memetakan Critical Thinking dan Practical Skill Peserta Didik SMP. Jurnal Inovasi Pendidikan IPA, 2(2), 213. https://doi.org/10.21831/jipi.v2i2.11963

Kristiyanto, S., Ashadi, A., Yamtinah, S., & Mulyani, S. (2019). Development of Computerized Testlet to Measure Science Process Skill on Stoichiometry. JKPK (Jurnal Kimia Dan Pendidikan Kimia), 4(3), 216. https://doi.org/10.20961/jkpk.v4i3.35198

Liu, H., Chen, X., & Liu, X. (2022). Factors influencing secondary school students' reading literacy: An analysis based on XGBoost and SHAP methods. Frontiers in Psychology, 13(September). https://doi.org/10.3389/fpsyg.2022.948612

OECD. (2019). PISA 2018 Assessment and Analytical Framework. In OECD Publishing.

Pinto, M., & Leite, C. (2020). Digital technologies in Support of Students Learning in Higher Education: Literature review. Digital Education Review, 37, 343–360. https://doi.org/10.1344/DER.2020.37.343-360

Retnawati, H. (2016). Proving Content Validity of Self-Regulated Learning Scale (The Comparison of Aiken Index and Expanded Gregory Index). Research and Evaluation in Education, 2(2), 155–164. https://doi.org/doi:http://dx.doi.org/10.21831/reid.v2i2.11029

Shrotryia, V. K., & Dhanda, U. (2019). Content Validity of Assessment Instrument for Employee Engagement. SAGE Open, 9(1). https://doi.org/10.1177/2158244018821751

Sudaryono, Rahardja, U., Aini, Q., Isma Graha, Y., & Lutfiani, N. (2019). Validity of Test Instruments. Journal of Physics: Conference Series, 1364(1). https://doi.org/10.1088/1742-6596/1364/1/012050