Chemined 13 (1) (2024)

Chemistry in Education

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

Development of Test Instruments Based on Critical Thinking Skills for Measuring Students' Minimum Competency in Thermochemical Material

Achdah Rachmatul Bariyah, Endang Susilaningsih, Harjono, and Woro Sumarni

Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia

Info Artikel

Accepted: 04-05-2024
Approved: 11-21-2024
Published: 12-27-2024

Keywords:
Minimum competency; critical thinking skills; thermochemistry

Abstract

The changes in 21st century life force students to quickly adapt in the field of education, especially in critical thinking skills and minimum competencies. This study aims to develop a test instrument based on critical thinking skills to analyze the minimum competency level of students. The research method is Research and Development with a 4D development design (Define, Design, Develop, and Disseminate). The subjects of the study were students of class XI MIPA SMA Muhammadiyah 1 Temanggung. Data collection methods consist of observation and testing of test instrument designs. Data collection instruments use interview sheets, documentation, test instruments, expert validation sheets, and questionnaire response sheets. Test and non-test data analysis techniques are analysis of the feasibility of test instruments, validity and reliability estimation, measurement of minimum competency levels, and analysis of student questionnaire responses. The results of the observation stated that teachers had implemented the Minimum Competency Assessment test but it was not optimal. The results of the instrument feasibility study were valid based on expert validation. The validity of the questions contained one misfit question. The reliability of the questions tended to be good. The results of the competency level analysis obtained 22% proficient, 26% capable, 28% basic, and 24% need special intervention. The results of the student questionnaire responses stated that they agreed that the instrument was feasible and successful to use. The conclusion of the instrument from the development was tested for feasibility, content validity, and a positive response was obtained.

Alamat korespondensi:
Building D6 Sekaran Campus, Gunungpati, (024)8508112, Semarang 50229
E-mail: achdahr56@gmail.com

P-ISSN 2252-6609
E-ISSN 2502-6852

INTRODUCTION

The changing times to the 21st century require learners to be able to adapt quickly and require the ability to process information correctly. The ability that can be developed by learners is the ability to think critically. Critical thinking skills is the ability to analyze information, evaluate situations, and make informed decisions. This ability makes learners look at a problem from different perspectives and think of ways to solve problems appropriately and effectively. Through critical thinking skills students learn how to formulate opinions, draw conclusions systematically, and be careful in making decisions (Is' ad & Sukarmin, 2022). Learners' ability to think can be explored by assessing the surrounding phenomena, then assessing through the learner's point of view (Rosnaeni, 2021). Indicators of critical thinking according to Alsaleh (2020) include 4 important aspects, namely identification of problems, analysis of important arguments (pros and cons), drawing conclusions,

and analyzing and evaluating the results of conclusions. These thinking skills can be measured using the Minimum Competency Assessment (Narassati et al., 2021).

The Minimum Competency Assessment (AKM) is essential for students to develop their skills and function positively in society (Pusmenjar, 2020). AKM measures two fundamental competencies: reading literacy and numeracy.The competencies assessed in reading literacy and numeracy include logical and systematic thinking skills, reasoning abilities using learned concepts and knowledge, and the ability to process and evaluate obtained information (Mustagfiroh, 2020). AKM is conducted to assess students' learning outcomes after completing a specific topic taught by the teacher. One of the subjects that can be evaluated using AKM is chemistry learning.

Chemistry learning includes both theoretical studies and laboratory research. Chemistry is the science that explores substances, including their composition, structure, properties, changes, and energetics, requiring both skills and reasoning (Arumsari et al., 2016). To this day, chemistry remains a challenging subject for students due to its complex and abstract nature, requiring extra effort to fully understand (Aini et al., 2024). Thermochemistry is one of the most difficult topics for students to grasp, as it involves understanding the concepts of heat changes and chemical reactions based on factual data (Annafi & Sri Mulyani, 2015).

Thermochemistry is closely related to the Minimum Competency Assessment (AKM) aspects of reading literacy and numeracy. Therefore, thermochemistry was chosen as the subject matter for developing a test instrument to measure students' minimum competencies.

METHODS

The development of the test instrument in this study is based on critical thinking skills in thermochemistry. The research was conducted at SMA Muhammadiyah 1 Temanggung, located on Jalan K. H. Samanhudi 6-8, Temanggung I, Temanggung District, Central Java. The research subjects were grade XI MIPA students at SMA Muhammadiyah 1 Temanggung, consisting of 20 students from XI MIPA 1 for the small-scale trial, 30 students from XI MIPA 2 for the large-scale trial, and 50 students from XI MIPA 1 and XI MIPA 3 for the implementation stage.

This study employs a quantitative approach using Research and Development (R&D) with the 4D model (Define, Design, Develop, Disseminate) developed by Thiagarajan (1974). The research process consists of four stages. The define stage involves analyzing previous research and conducting a literature review. The design stage focuses on developing the test instrument and questionnaire. The develop stage includes testing the instrument through small-scale, large-scale, and implementation trials. Finally, the dissemination stage is conducted by publishing the research findings in academic articles.

Data collection was conducted through four methods: (1) observation, (2) testing of the Minimum Competency Assessment (AKM)-based test instrument, (3) distribution of questionnaires, and (4) documentation. The data analysis process involved content validity analysis

and Rasch modeling for test instruments. Meanwhile, for non-test instruments, the analysis included content validity and response reliability provided by students..

RESULT AND DISCUSSION

Validity of AKM Test Instrument

Content validity was conducted on the test instrument to assess its feasibility. The validation process was carried out by three experts in the field, consisting of two chemistry lecturers and one chemistry teacher. The obtained scores can be seen in Table 1.

Table 1. AKM Test Instrument Validity Score

Validator	Total Score	Category	Description
Validator 1	24	Valid	Can be used with revision
Validator 2	28	Valid	Can be used without revision
Validator 3	29	Valid	Can be used without revision

Based on the results obtained, it was concluded that the test instrument was deemed valid by all three validators. The average validation score was 27 out of a total score of 32, placing it in the valid category. The validation results indicate that the developed test instrument is suitable for trial implementation after minor revisions, such as improving the clarity of question wording, adjusting image placement, and refining answer choices.

The expert validation results indicate that the test instrument is suitable for trial implementation. However, the small-scale testing process faced challenges due to limited time and students' lack of readiness. As a result, students experienced difficulties in answering the test questions comprehensively.

Estimated Reliability of the Test Instrument

The reliability assessment of the test instrument was conducted using the Rasch model to obtain values for person reliability, item reliability, and Cronbach's Alpha. The results of the reliability analysis are presented below.

Table 2. AKM Test Instrument Reliability Score

Analysis	Test Stages			
	Small Scale		Large Scale	
	Dicothomy	Polytomy	Dichotomy	Polytomi
Person Realibility	0,81 (Good)	0,75 (fair)	0,71 (Fair)	0,78 (Fair)
Item Realibility	0,79 (Fair)	0,50 (Weak)	0,69 (Fair)	0,91 (Very Good)
Cronbach Alpha	0,84 (Very Good)	0,78 (Good)	0,74 (Good)	0,75 (Good)

Based on the data from Table 2, it can be concluded that in each testing phase, the *person reliability* value shows an improving trend (>0.67, categorized as good). This indicates that students' response consistency is fairly good (Purniasari et al., 2021). Each testing phase also demonstrates an increasing *item reliability* value (>0.67, categorized as good), which reflects a high level of reliability in the test items. Additionally, the *Cronbach's Alpha* reliability value shows continuous improvement, meaning that the overall interaction between students and test items is becoming more reliable.

Item Fit Order

Item validity is used to analyze the accuracy of the test items developed and to determine whether revisions are necessary. The results of the item validity analysis for each testing phase are presented in Table 3.

Table 3. Recapitulation of AKM Test Instrument Validity

Analysis	Test Stages	
	Small Scale	**Large Scale**
1	Valid	Valid
2	Valid	Valid
3	Valid	Valid
4	Valid	Valid
5a	Valid	Valid
5b	Valid	Valid
5c	Valid	Valid
6	Valid	Valid
7	Valid	Valid
8	Valid	Valid
9	Valid	Valid
10	Valid	Valid
11	Valid	Valid
12a	Valid	Valid
12b	Not Valid	Valid
12c	Valid	Valid
13a	Valid	Valid
13b	Valid	Valid
13c	Valid	Valid
14	Valid	Valid
15a	Valid	Valid
15b	Valid	Valid
16	Valid	Valid

17a	Valid	Valid
17b	Valid	Valid
17c	Valid	Valid
18	Valid	Valid
19a	Valid	Valid
19b	Valid	Valid
20	Valid	Valid

Based on Table 3, there is no consistent pattern of item invalidity across all testing phases. Some test items that were initially invalid during the small-scale trial became valid in the large-scale trial, indicating that the revisions made had a positive impact. Conversely, test items that were valid in the small-scale trial but later became invalid in the large-scale trial may have been influenced by several factors, such as differences in the number of students and variations in students' ability levels during testing (Mahmudah et al., 2016).

Item Measure

The Item Measure analysis is used to determine the difficulty level of each test item. In this study, the difficulty level was identified based on the logit values obtained in the measure column. Based on the analysis conducted during the implementation phase using dichotomous items, a standard deviation of 0.98SD was obtained, classifying the test items into four categories: very difficult, difficult, moderate, and easy. The very difficult category, with a logit value greater than 0.98SD, includes items 15b, 10, 17c, 19a, and 15a. The difficult category, with logit values ranging from 0.0 to 0.98SD, includes items 12a, 19b, 11, 12b, 13a, 14, 13b, 17a, 20, and 1. The moderate category, with logit values ranging from 0.0 to -0.98SD, includes items 16, 3, and 5c, while the easy category, with a logit value less than -0.98SD, includes items 13c, 17b, 2, 5a, 5b, 7, and 12c.

For the implementation analysis using polytomous items, a standard deviation of 0.56SD was obtained, classifying the items into three categories: very difficult, difficult, and moderate. The very difficult category, with a logit value greater than 0.56SD, includes item 9. The difficult category, with logit values ranging from 0.0 to 0.56SD, includes item 8. Meanwhile, the moderate category, with logit values ranging from 0.0 to -0.56SD, includes items 6, 18, and 4.

Profile of Students' Minimum Competency Achievement

The analysis of students' minimum competencies aims to assess their basic abilities in *reading literacy* and *numeracy* within the *thermochemistry* topic. Students' minimum competencies are determined based on their test results, which are then analyzed using *Microsoft Excel* and interpreted into four different competency levels: (1) *Needs Special Intervention*, (2) *Basic*, (3) *Proficient*, and (4) *Advanced* (Pusmenjar, 2020). The classification of students' minimum competency levels is based on the *standard deviation (SD)* and the *average score* of the students.

The results of the minimum competency level assessment using *Microsoft Excel* are presented in Figure 1.

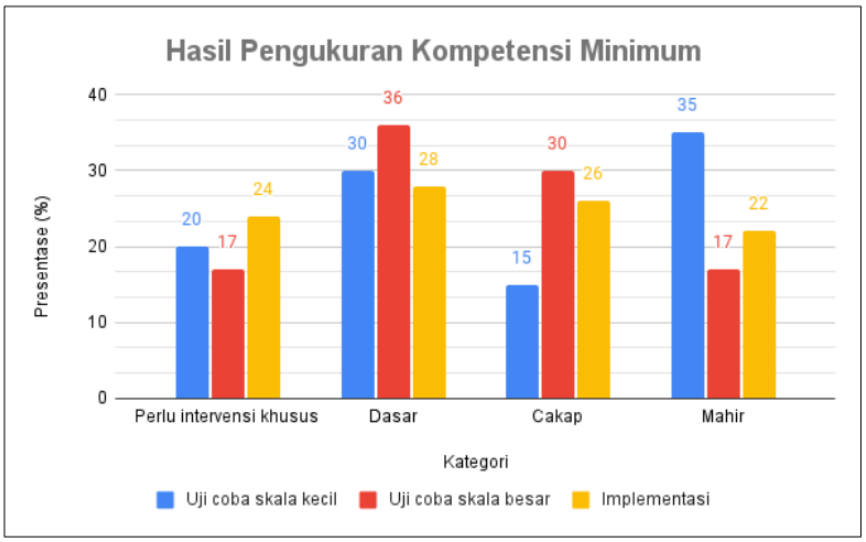

Figure 1. Results of Students' Minimum Competency Level Assessment

Based on Figure 1, the analysis of students' minimum competency levels varies across the small-scale trial, large-scale trial, and implementation phases. This difference is due to the different subjects involved in each testing stage. The analysis results show that in the large-scale trial and implementation phases, students' minimum competency levels were predominantly at the *basic* level, with percentages of 36% and 28%, respectively. Meanwhile, the small-scale trial was dominated by students at the *advanced* level, with a percentage of 35%. This criterion can be considered *fairly good* since students have demonstrated a sufficient level of understanding and problem-solving skills in responding to the *Minimum Competency Assessment (AKM)* test instrument.

The factors contributing to students not achieving the minimum competency level or being at the lowest level can be categorized into *internal* and *external* factors. Internal factors include students' lack of attentiveness when reading test items, low reading motivation, and weak understanding of prerequisite materials. External factors include negative peer influence during test completion and ambiguously worded questions that lead to multiple interpretations, causing confusion among students (Priliyanti et al., 2021).

Analysis of Students' Responses to the Developed AKM Test Instrument

The analysis of students' responses was conducted using data obtained from questionnaires distributed after completing the *AKM test instrument*. This response analysis aims to assess the effectiveness of the developed test instrument. The reliability of the questionnaire was found to be *0.88* in the small-scale trial, *0.89* in the large-scale trial, and *0.85* in the implementation phase. These reliability results indicate that the questionnaire is considered *reliable.*

CONCLUSION

The AKM test instrument based on critical thinking skills in thermochemistry was found to be valid and reliable, confirming its feasibility for use. The overall profile of students' minimum competency levels indicates that 22% of students reached the advanced level, 26% were at the proficient level, 28% were at the basic level, and 24% required special intervention. Students' responses to the developed instrument were positive, with 89% of students agreeing that the test instrument was effective, as shown in the response questionnaire summary.

REFERENCE

Aini, D. A., Solekha, N. A., Juvitasari, S., Nisfah, U. A., Tia, R., Siti Aisyah, R. S., Irhamni, I., & Zidny, R. (2024). Survei Analisis Kebutuhan Pengembangan Lembar Kerja Peserta Didik Berbasis Socio Scientific Issues pada Materi Asam Basa. Jambura Journal of Educational Chemistry, 6(1), 1–10. https://doi.org/10.37905/jjec.v6i1.23505

Alsaleh, N. J. (2020). Teaching Critical Thinking Skills: Literature Review. Turkish Online Journal of Educational Technology-TOJET, 19(1), 21–39.

Annafi, N., & Sri Mulyani, D. (2015). Pengembangan Lembar Kegiatan Peserta Didik Berbasis Inkuiri Terbimbing Pada Materi Termokimia Kelas Xi Sma/Ma. 4(3), 21–28.

Arumsari, L. T., Rosilawati, I., & Kadaritna, N. (2016). Pengembangan Instrumen Asesmen Keterampilan Proses Sains Pada Materi Teori Tumbukan. Jurnal Pendidikan Dan Pembelajaran Kimia, 5(1), 140–151.

Is' ad, N., & Sukarmin, S. (2022). Implementation of Problem-Solving Learning Model Assisted by Student Worksheets to Improve Critical Thinking Skills in The Context of Reaction Rate. Jurnal Pijar Mipa, 17(2), 199–208.

Mahmudah, R., Pramudya, Y., & Sulisworo, D. (2016). Analisis Validitas Butir Soal Certainty of Respons Index (CRI) untuk Identifikasi Miskonsepsi Materi Tata Surya dan Fenomena Astronomi. Prosiding Seminar Nasional Pendidikan Sains, 579–587.

Mustagfiroh. (2020). Memanfaatkan Hasil Asesmen Kompetensi Minimum (AKM) untuk Mendesain Multi Modal Learning. Jurnal Guru Inovatif, 2, 48–62.

Narassati, N. A., Saleh, R., & Arthur, R. (2021). Pengembangan alat evaluasi berbasis hots menggunakan aplikasi quizizz pada mata pelajaran mekanika teknik dalam pembelajaran jarak jauh. Jurnal Pendidikan Teknik Sipil, 3(2), 169–180.

Priliyanti, A., Muderawan, I. W., & Maryam, S. (2021). Analisis kesulitan belajar siswa dalam mempelajari kimia kelas XI. Jurnal Pendidikan Kimia Undiksha, 5(1), 11–18.

Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis butir soal ujian sekolah mata pelajaran kimia SMA N 1 kutowinangun tahun pelajaran 2019/2020 menggunakan model iteman dan rasch. Jurnal Pendidikan Kimia, 10(2), 205–214.

Pusmenjar. (2020a). AKM dan Implikasinya pada Pembelajaran. In Pusat Assessment dan Pembelajaran, Jakarta.

Pusmenjar. (2020b). Desain Pengembangan Soal Asesmen Kompetensi Minimum. Rosnaeni, R. (2021). Karakteristik dan Asesmen Pembelajaran Abad 21. Jurnal Basicedu, 5(5), 4334–4339.