Proceeding of 2nd International Conference on Education, Society and Humanity
e-ISSN : 2986-5832
p-ISSN : 2986-6979
Vol. 03 No. 01 (2025)
Available online at https://ejournal.unuja.ac.id/index.php/icesh

CONTENT VALIDITY AND RELIABILITY OF A STUDENT PERFORMANCE ASSESSMENT INSTRUMENT IN COLLABORATIVE PROBLEM SOLVING-BASED MATHEMATICS LEARNING

Isnurani[1], Wardani Rahayu[2], Iva Sarifah[3]
[1]Universitas Pamulang, Tangerang Selatan, Banten, Indonesia
[2,3]Universitas Negeri Jakarta, Indonesia
E-mail: dosen01193@unpam.ac.id[1],wardani.rahayu@unj.ac.id[2],ivasarifah@unj.ac.id[3]

Abstract:

This study aimed to examine the content validity and internal reliability of a student performance assessment instrument for mathematics learning based on Collaborative Problem Solving (CPS). The participants of this study were 45 students from a private high school in South Tangerang, Banten. The instrument is a self-assessment questionnaire with 32 statements based on four dimensions of Collaborative Problem Solving (CPS), namely cognitive, social, emotional, and metacognitive. Content validity in this study uses Aiken's V index test, while internal reliability is analyzed using Cronbach's Alpha. The results showed that most items had fairly good validity with an Aiken's V value ≥ 0.60, but there were 11 items with values below 0.60 that needed improvement. Meanwhile, the reliability value of the instrument was very high with a Cronbach's Alpha value above 0.90 for each dimension and 0.986 overall. These findings confirm that the instrument is suitable for use as a comprehensive measurement tool in supporting mathematics learning transformation based on collaborative problem solving.

Keywords: *Performance Assessment, Collaborative Problem Solving, Mathematics Learning*

INTRODUCTION

Mathematics plays a crucial role in developing critical thinking, analytical, and problem solving skills. Mathematics learning requires students to develop logical, systematic, and strategic thinking patterns that can be applied in real life (Griffin & Care, 2015; Kapur, 2014). Therefore, mathematics is the main vehicle for training students' higher order thinking skills (HOTS).

However, Indonesian students' achievements in international assessments are still not encouraging. The 2022 PISA report released by the OECD (2023) confirms that Indonesian students have difficulty solving problems that require critical reasoning and complex problem solving. These findings reflect the urgent need to improve the quality of mathematics learning in schools.

One of the causes of these low achievements is the conventional learning pattern that is still dominant, in which teachers place more emphasis on the final result than on the students' thinking process. This approach does not provide enough space for students to discuss, interact, and develop collaborative skills. In fact, collaboration is an important part of 21st-century competencies that are essential in facing global challenges (Wardani, 2020).

In response to this need, Collaborative Problem Solving (CPS) has emerged as a learning model that emphasizes problem solving through cooperation. CPS encourages students to share ideas, develop strategies, and make decisions

together. Collaborative Problem Solving has been shown to not only improve conceptual understanding of mathematics but also build social skills such as communication, coordination, and conflict resolution (Fiore et al., 2017; Griffin, 2010). Thus, Collaborative Problem Solving can be a relevant learning innovation to improve the quality of mathematics education.

Despite its potential, the application of Collaborative Problem Solving still faces challenges in terms of assessment. Existing assessment instruments tend to be limited to cognitive achievements and final results, while the process of interaction, decision making, and metacognitive reflection are rarely evaluated (Widodo & Wahyudi, 2019). In fact, the success of Collaborative Problem Solving-based learning is not only measured by the final answer, but also by the dynamics of students in collaborating and strategizing (Roschelle et al., 2019).

Comprehensive student performance assessment instruments need to be developed to evaluate various dimensions in Collaborative Problem Solving. Fiore et al. (2017) state that Collaborative Problem Solving-based performance assessments should at least cover three main aspects: building shared understanding, making the right decisions, and coordinating teamwork. Furthermore, Saadati and Felmer (2021) emphasize the need for a balanced assessment of cognitive, social, emotional, and metacognitive dimensions to reflect the complexity of collaborative learning in mathematics.

This effort aligns with the direction of national education policy. The Merdeka Curriculum launched by the Ministry of Education, Culture, Research, and Technology (2022) emphasizes the importance of authentic assessment that not only assesses the final results but also the learning process of students. Therefore, the existence of a Collaborative Problem Solving-based performance assessment instrument can be seen as an important innovation in supporting the transformation of mathematics learning oriented towards strengthening 21st-century skills.

Several previous studies have developed CPS based assessment instruments (Medina et al., 2019; Wardani, 2020). However, the instruments developed emphasize cognitive and social aspects, while the emotional and metacognitive dimensions are still under-explored. This shortcoming indicates the need for research that develops a more comprehensive assessment of student performance. The instrument used in this study was developed and its construct validated in a prior stage of this research. Building on that foundation, the present study proceeds to the try-out stage of instrument testing, using data from a sample of 45 respondents drawn from the broader field try out phase. Therefore, this study focuses on Collaborative Problem Solving based student performance assessment as an innovation in mathematics learning in terms of content validity and internal reliability of the instrument.

RESEARCH METHODS

This study used a development research method with a model adapted from Mardapi's (2017) development stages with research procedures designed to produce a Collaborative Problem Solving-based student performance assessment instrument that not only assesses the final results but also the students' collaboration process in solving mathematical problems. The development procedures used in this study included: (1) Planning, (2) Development, (3) Content validity, and (4) Field testing (initial). The participants of this study were 45 high school students in South Tangerang. The instrument was a self-assessment questionnaire with 32 items compiled based on four dimensions of CPS. This study used a five-point Likert scale. Data analysis was conducted in two stages. First, content validity was assessed using Aiken's V index to assess the representativeness and suitability of the items to the indicators. Second, internal reliability was analyzed using Cronbach's Alpha to measure the consistency between items in each dimension. The instrument was declared valid if the Aiken's V value was ≥ 0.60 and reliable if the Cronbach's Alpha value was ≥ 0.70, while an α value of ≥ 0.90 indicated a highly reliable category (Azwar, 2019;

Widoyoko, 2018).

DISCUSSION

Content validity in this study was used to measure the extent to which the instrument items were appropriate for the construct to be measured. Content validity was measured based on Aiken's V value obtained from the distribution of student responses.

Table 1. Aiken's V Index

No	Dimension	Average Score	Aiken's V	Criteria
1	Cognitive	3.77	0.69	Valid enough
2		3.30	0.58	Needs revision
3		3.33	0.58	Needs revision
4		3.19	0.55	Needs revision
5		3.38	0.60	Valid enough
6		3.25	0.56	Needs revision
7		3.71	0.68	Valid enough
8		3.22	0.56	Needs revision
9	Social	3.80	0.70	Valid enough
10		3.84	0.71	Valid enough
11		3.60	0.65	Valid enough
12		3.73	0.68	Valid enough
13		3.55	0.64	Valid enough
14		3.82	0.70	Valid enough
15		3.66	0.67	Valid enough
16		3.28	0.57	Needs revision
17	Emotional	3.89	0.72	Valid enough
18		3.91	0.73	Valid enough
19		3.86	0.72	Valid enough
20		3.82	0.70	Valid enough
21		3.22	0.56	Needs revision
22		3.95	0.74	Valid enough
23		3.18	0.55	Needs revision
24		3.15	0.54	Needs revision
25		3.19	0.55	Needs revision
26	Metacognitive	3.82	0.70	Valid enough
27		3.77	0.69	Valid enough
28		3.64	0.66	Valid enough
29		3.78	0.69	Valid enough
30		3.58	0.65	Valid enough
31		3.67	0.67	Valid enough
32		3.81	0.70	Valid enough

Based on Table 1, it indicates that most items have a value of ≥ 0.60, so they can be considered sufficiently valid. However, there are 11 items with a value of < 0.60, requiring editorial improvements to make them more operational. This finding indicates that, in general, the instrument is sufficiently valid for the initial trial phase, and revisions to the 11 items will improve the clarity of the items before further trials are conducted.

Reliability indicates the internal consistency of instrument items in measuring the same construct. Cronbach's Alpha is used because it is suitable for the Likert scale.

Table 2. Cronbach's Alpha Values per Dimension and Total

No	Dimension	Total Items	Alpha	Interpretation
1	Cognitive	8	0.939	Highly reliable
2	Social	8	0.969	Highly reliable
3	Emotional	8	0.935	Highly reliable
4	Metacognitive	8	0.955	Highly reliable
	Total	32	0.986	Highly reliable

Based on Table 2, the total Alpha (α) value of 0.986 indicates that the instrument has very high internal consistency. This means that students' responses to each item of the instrument show a stable and homogeneous pattern. This high reliability indicates that the items in each dimension consistently measure the same construct. This indicates that the instrument is suitable for assessing student performance in Collaborative Problem Solving-based mathematics learning and is ready for use in further trials after minor revisions to several items.

The results of preliminary tests show that the Collaborative Problem Solving (CPS) based student performance assessment instrument in mathematics learning has met the criteria for content validity and internal reliability. Based on the results of the analysis using Aiken's V index, a value between 0.54 and 0.74 was obtained. Most items showed values above 0.60, indicating that the statements adequately represented the construct being measured, while eleven items were still below 0.60 and needed improvement. According to Azwar (2019), items with an Aiken's V value ≥ 0.60 can be declared valid in terms of content, while values below that indicate that the statements are still too general, ambiguous, or less appropriate for the indicators.

These findings indicate that most of the instrument items are in line with the four main dimensions of Collaborative Problem Solving, namely cognitive, social, emotional, and metacognitive. The social and emotional dimensions show higher results than the cognitive and metacognitive dimensions. This indicates that respondents find it easier to understand statements related to cooperation, communication, and emotional control in collaborative activities. These findings are in line with Wardani (2020) opinion that students' collaborative skills can be effectively measured through indicators of communication, coordination, and group responsibility.

Furthermore, in the cognitive and metacognitive dimensions, several items with low validity scores were found. Revisions to the cognitive dimension were made because many errors arose due to statements that were too general, such as "I can explain how geometric transformations work." The suggested revision is to narrow the scope directly to the material indicators: "I can explain changes in flat shapes through translation, rotation, reflection, or dilation." In the social dimension, weak items tended to be related to collaborative attitudes, which should be clarified with more active verbs, such as "I actively listen to and respond to my friends' opinions in group discussions." According to Widoyoko (2018), clarity of wording and accuracy of operational verbs are very important to ensure consistency between statements and the indicators of ability being measured. Therefore, revisions to items with low Aiken's V values are directed toward the use of more specific verbs and mathematical contexts relevant to the

context of mathematics learning materials.

Furthermore, based on the Cronbach's Alpha calculation results, the four dimensions have values above 0.90, with a total value of 0.986. Based on the criteria proposed by Azwar (2019) and Widoyoko (2018), a reliability value above 0.90 is classified as highly reliable, indicating that each item in a dimension has high internal consistency in measuring the same construct. This result is also reinforced by Sugiyono (2017), who explains that high reliability indicates stability between items and consistency in measurement results when the instrument is reused under different conditions.

Specifically, the social dimension showed the highest reliability value (α = 0.969). This indicates that statements related to interaction, teamwork, and communication among students have very strong coherence. These results are in line with Griffin and Care's (2015) theory that collaborative success in problem solving depends on good social skills, such as building mutual understanding and adjusting strategies with peers in the group. Meanwhile, the metacognitive dimension (α = 0.955) also has high reliability, indicating that indicators of self-reflection and learning strategy control are consistently responded to by students. High reliability across all dimensions proves that this instrument is well structured and reliable for measuring student performance in the CPS context. Thus, this instrument is suitable for use as an initial measurement tool in mathematics learning. Several items that still need revision will be refined in the next development stage so that the instrument is more representative of the construct being measured. With continuous improvement, this instrument is expected to become an assessment innovation that supports the transformation of mathematics learning towards strengthening collaborative skills and 21st-century problem solving.

CONCLUSION

The Collaborative Problem Solving (CPS) based student performance assessment instrument in mathematics learning has fairly good content validity, although there are 11 items with Aiken's V values below 0.60 that need to be revised to be clearer and more specific. The reliability of the instrument is very high, with Cronbach's Alpha values above 0.90 in each dimension and 0.986 overall, indicating that the instrument is consistent in measuring student skills. Thus, this instrument is suitable for use as a comprehensive measuring tool.

REFERENCES

Azwar, S. (2019). Reliabilitas dan validitas. Yogyakarta: Pustaka Pelajar.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., & Veloso, M. (2017). Collaborative problem solving: Considerations for the National Assessment of Educational Progress. Journal of Educational Psychology, 109(8), 1049–1067.

Griffin, P., & Care, E. (2015). Assessment and teaching of 21st century skills: Methods and approach. Springer.

Kapur, M. (2014). Productive failure in learning mathematics. Instructional Science, 42(4), 523–538.

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2022). Kurikulum Merdeka. Jakarta: Kemendikbudristek.

Mardapi, D. (2017). Pengukuran, penilaian, dan evaluasi pendidikan. Yogyakarta: Parama Publishing.

Medina, M. J., Martinez, A., & Ortega, C. (2019). Collaborative learning and

assessment: A model for the 21st century. International Journal of Learning, Teaching and Educational Research, 18(8), 158–174.
OECD. (2023). PISA 2022 results: Learning during (and from) disruption. OECD Publishing.
Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. (2019). Collaborative problem solving in mathematics: Supporting and measuring teachers and students in classrooms. Journal of Educational Psychology, 111(3), 470–492.
Saadati, F., & Felmer, E. (2021). Assessing collaborative problem solving in mathematics: A framework. Educational Studies in Mathematics, 106(1), 1–20.
Sugiyono. (2017). Metode penelitian kuantitatif, kualitatif, dan R&D. Bandung: Alfabeta.
Wardani, S. (2020). Asesmen keterampilan kolaboratif siswa melalui pembelajaran matematika. Jurnal Riset Pendidikan Matematika, 7(2), 165–178.
Widodo, S. A., & Wahyudi, W. (2019). Analisis asesmen autentik dalam pembelajaran matematika. Infinity Journal, 8(1), 45–58.
Widoyoko, E. P. (2018). Evaluasi program pembelajaran. Yogyakarta: Pustaka Pelajar.