Jurnal Pemikiran dan Penelitian Pendidikan Matematika (JP3M)
Journal homepage: https://journal.rekarta.co.id/index.php/jp3m/index
ISSN 2622-7673 (Online) | ISSN 2622-8246 (Print) | Vol. 9, No. 2, November 2026
https://doi.org/10.36765/jp3m.v9i2.993

From Bloom to SOLO: Integrating Critical Thinking and Sustainability Context in Geometry Assessment Instruments for Phase D Students

Fatimah Leila Danah Luthfiyah[1], Abi Suwito[2], Erfan Yudianto[3], Arif Fatahillah[4]
[1,2,3,4]Mathematics Education Study Program, Master's Program, Universitas Jember, Indonesia.
✉ Author Corresponding: fatimahleila02@gmail.com

ABSTRACT

Quality geometry assessment must be able to measure students' higher-order thinking skills gradually. This article presents a synthesis of two research phases: (1) the development of a summative assessment instrument based on Bloom's Taxonomy to measure critical thinking in quadrilateral material, and (2) a conceptual framework for formative and summative assessment instruments based on SOLO Taxonomy with a waste management context in cuboid material. The first study employed a 4D model with 56 Phase D students, producing an instrument that was valid ($Va = 3.79$), reliable ($r = 0.95$), practical ($86 - 89\%$), and effective ($55 - 77\%$). Findings indicated that students excelled in interpretation but were weak in evaluation and inference. Based on these findings, this article proposes a SOLO-based assessment framework that offers hierarchical classification of student responses and integrates Education for Sustainable Development (ESD) through a waste management context. This framework is expected to address the limitations of Bloom-based assessment in measuring depth of understanding and providing more specific feedback.

Keywords: SOLO taxonomy; critical thinking; geometry assessment; sustainability context, Phase D

Article History:
Received: 09-10-2026
Revised : 23-10-2026
Accepted: 29-10-2026
Online : 15-11-2026

How to Cite (APA style):
Luthfiyah, F. L. D., et.al (2026). From Bloom to SOLO: Integrating Critical Thinking and Sustainability Context in Geometry Assessment Instruments for Phase D Students. *Jurnal Pemikiran dan Penelitian Pendidikan Matematika (JP3M)*, 9(2), 49-57. https://doi.org/10.36765/jp3m.v9i2.993

This is an open access article under the CC–BY-SA license

1. INTRODUCTION

Assessment plays a strategic role in improving the quality of mathematics learning as it serves as a tool for mapping students' understanding and providing feedback for learning improvement (Toharudin et al., 2025). The quality of assessment instruments largely determines the accuracy of information obtained about student learning progress, making the development of valid and reliable instruments an urgent necessity (Audiwinanda & Mahmudi, 2025). Empirical studies have consistently shown that assessment practices in schools often remain limited to measuring low-level cognitive abilities, such as remembering and understanding, and have not yet reached the stages of analyzing, evaluating, and creating (Evangelista & Rabut, 2025; Zana et al., 2024). This condition creates a gap between curriculum demands that emphasize 21st-century skills and the reality of assessment in the field (Zana et al., 2024). Therefore, a breakthrough is needed in developing assessment instruments capable of measuring higher-order thinking skills more comprehensively (Kania & Kyaruzi, 2026).

One theoretical framework that has long been used to measure higher-order thinking skills is Bloom's Taxonomy, which encompasses cognitive process dimensions from remembering (C1) to creating (C6) (Toharudin et al., 2025). The urgency of this research is rooted in the need to develop a more nuanced assessment tool for Phase D students (junior high school level). Phase D is a critical transition period where students are expected to master abstract and complex mathematical concepts, yet their thinking skills are often assessed using instruments that fail to capture the depth of their understanding. Previous research has developed a summative assessment instrument based on Bloom's Taxonomy to measure critical thinking skills in quadrilateral material for Phase D students. The instrument was designed by integrating six critical thinking indicators, namely interpretation, analysis, evaluation, inference, explanation, and self-regulation Facione (2015), into contextual essay questions. It is important to note that these indicators, while rooted in the social context of critical thinking, were operationalized for mathematics education by aligning them with cognitive processes required for solving geometric problems, such as interpreting information, analyzing relationships, and evaluating solutions. The results showed that the instrument was valid with an average score of 3.79, reliable with a Cronbach's Alpha coefficient of 0.95, practical with a percentage of 86 – 89%, and effective with critical thinking achievement of 55 – 77% (Priyanto et al., 2024). However, the research findings also revealed that students tended to excel in interpretation compared to evaluation and inference, indicating the need for an assessment framework that can better classify students' depth of understanding gradually (Yuanita et al., 2023).

SOLO (Structure of Observed Learning Outcomes) Taxonomy offers a more nuanced framework as it classifies the quality of student responses into five hierarchical levels: prestructural, unistructural, multistructural, relational, and extended abstract (Adeniji et al., 2022). Unlike Bloom's Taxonomy which is oriented towards learning objectives, SOLO Taxonomy specifically observes the structure of student answers to determine the extent to which they can integrate information and build relationships between concepts (Evangelista & Rabut, 2025). The main advantage of SOLO Taxonomy lies in the existence of a prestructural level that helps educators identify misconceptions or student ignorance as a starting point for learning (Ariani et al., 2026). Previous research has proven that SOLO-based instruments are able to measure students' thinking levels more accurately and encourage movement from multistructural to relational and extended abstract levels (Svensäter & Rohlin, 2023). Therefore, SOLO Taxonomy is considered relevant as a framework for developing assessment instruments capable of addressing the limitations of Bloom's Taxonomy in measuring depth of understanding (Fernandez & Guzon, 2025).

The integration of sustainability context in mathematics assessment is becoming increasingly important along with growing global awareness of environmental issues, particularly waste management (Giannini, 2020). Education for Sustainable Development (ESD) emphasizes the need to integrate sustainability principles into various curriculum areas so that students can connect knowledge with real-world challenges (Purnamasari et al., 2026). The waste management context offers an opportunity to link geometry concepts, such as volume and surface area of cuboids, with calculations of container capacity and material efficiency for making recycled waste bins (Sembiring et al., 2026). Several studies have shown that the integration of ESD in mathematics learning not only improves conceptual understanding but also fosters environmental awareness and students' social responsibility (Lafuente-Lechuga et al., 2024; Vásquez et al., 2023). However, there is still a research gap in integrating ESD into SOLO Taxonomy-based mathematics assessment instruments on cuboid material for Phase D students.

Previous studies have not specifically addressed the development of a comprehensive assessment framework that combines the hierarchical structure of SOLO Taxonomy with the contextual relevance of ESD for this specific mathematical topic and educational level (Andikos et al., 2025; Ramadhani et al., 2025).

Based on the background and research gap that have been elaborated, this article aims to present a synthesis of two interrelated research phases. The first phase has produced a summative assessment instrument based on Bloom's Taxonomy to measure critical thinking skills in quadrilateral material, with empirical findings on its validity, reliability, practicality, and effectiveness (Priyanto et al., 2024). The second phase offers a conceptual framework for formative and summative assessment instruments based on SOLO Taxonomy with a waste management context in cuboid material (Svensäter & Rohlin, 2023). This synthesis is expected to provide theoretical and practical contributions to the development of more meaningful geometry assessment, as well as serve as a reference for researchers and education practitioners in designing assessment instruments that integrate higher-order thinking skills and sustainability issues (Dao et al., 2026).

2. METHODS

This article employs a qualitative synthesis approach to two interrelated research phases, focusing on thematic analysis of empirical findings and the proposed conceptual framework (Creswell, 2012). The research procedure is systematically structured into two distinct phases, which are explained as follows:

1) The first research phase employed a 4D development model (Define, Design, Develop, Disseminate) involving Phase D students at SMPN 1 Rambipuji. The Define stage involved a needs analysis through interviews with mathematics teachers to identify the need for critical thinking instruments in quadrilateral material. The Design stage focused on drafting test grids, essay questions integrating six critical thinking indicators from Facione (2015) and cognitive levels of Bloom's Taxonomy C4-C6, and a scoring rubric. The Develop stage included expert validation (average score 3.79) and small group trials (n=10) to assess reliability, item difficulty, and discrimination. The Disseminate stage involved distributing the instrument to mathematics teachers and uploading it to online platforms. A large group trial was then conducted with 56 students at SMPN 1 Rambipuji, divided into two groups (26 and 30 students), both receiving the same treatment. Data collection in the first phase was conducted through expert validation, small group trials, and large group trials, with data analysis including validity, reliability, item difficulty, item discrimination, practicality, and effectiveness (Hobri et al., 2021).
2) The second phase is a conceptual framework designed based on needs analysis and literature review, employing the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) and integrating five levels of SOLO Taxonomy and waste management context as a manifestation of ESD (Fadhli et al., 2023; Maruf et al., 2024). The development procedure in the second phase is systematically designed through five cyclical ADDIE stages (Pankaew et al., 2025). The analysis stage involves identifying learning needs, characteristics of Phase D students, and analyzing cuboid material in the Merdeka Curriculum (Toharudin & Dewi, 2025), as well as preliminary observations at two schools indicating the need for more in-depth assessment instruments. The design stage includes the preparation of test grids, essay questions, SOLO-based scoring rubrics, and formative assessment instruments in the form of self-assessment, peer assessment, and observation sheets (Ariani et al., 2026; Toharudin &

Dewi, 2025). The development stage involves expert validation, small group trials, and item analysis, while the implementation and evaluation stages are designed to test the validity and reliability of the instrument in actual learning situations (Asri et al., 2024). The developed instrument encompasses formative and summative aspects in an integrated manner, with a rubric adapting the SOLO transition levels from Fernandez & Guzon (2025) to assess students' conceptual understanding more nuancedly.

3. RESULT AND DISCUSSION

Findings from the first research phase indicate that the summative assessment instrument based on Bloom's Taxonomy developed meets the criteria of being valid, reliable, practical, and effective for measuring critical thinking skills of Phase D students in quadrilateral material (Priyanto et al., 2024). The operationalization of Bloom's Taxonomy and critical thinking indicators is detailed in Table 1, showing how each indicator was integrated into the instrument's questions.

Table 1. Operationalization of Bloom's Taxonomy and Critical Thinking Indicators in the Assessment Instrument

Bloom's Cognitive Level	Critical Thinking Indicator (Facione, 2015)	Operational Definition in the Assessment Instrument
C4 (Analyze)	Interpretation	Students are able to write down relevant information from the problem, explain what is known and asked, and identify key elements of the quadrilateral problem.
C4 (Analyze)	Analysis	Students are able to outline problem-solving steps based on quadrilateral concepts, connecting available information with relevant mathematical concepts.
C5 (Evaluate)	Evaluation	Students are able to re-examine the steps of solving the problem, evaluate the accuracy of calculations, and determine if the final result is in accordance with the problem.
C5 (Evaluate)	Inference	Students are able to draw conclusions based on a clear relationship between the information in the problem and state the results of the analysis with supporting explanations.
C6 (Create)	Explanation	Students are able to provide a detailed explanation of the final answer, including the reasons why the method or rules used to solve the problem were chosen.
C6 (Create)	Self-Regulation	Students are able to reflect on and correct their answers if errors are found, and review the overall consistency between the final result and the problem's question.

The summary of the first phase's empirical findings is presented in Table 2, detailing the instrument's quality criteria and effectiveness.

Table 2. Summary of Empirical Findings from the First Research Phase

Quality Criteria	Result	Category	Notes
Validity (Average Score)	3.79	Valid	Based on expert validation of content and language aspects.
Reliability (Cronbach's Alpha)	0.95	Very High	Indicates high consistency of the instrument.
Item Difficulty	30% Easy, 50% Medium, 20% Difficult	Ideal	The distribution is considered balanced.
Item Discrimination	8 out of 15 items in the "Good" category	Good	The items are able to distinguish between high and low-ability students.
Practicality (Group 1, n=26)	89%	Practical	Based on student response questionnaires.
Practicality (Group 2, n=30)	86%	Practical	Based on student response questionnaires.
Effectiveness (Group 1, n=26)	55%	Sufficient	Average critical thinking achievement.
Effectiveness (Group 2, n=30)	77%	High	Average critical thinking achievement.

Further analysis of critical thinking indicator achievement revealed a consistent pattern in both trial groups, as shown in Table 3.

Table 3. Achievement of Critical Thinking Indicators in Both Trial Groups

Critical Thinking Indicator	Group 1 (n=26) Achievement	Group 2 (n=30) Achievement	Category
Interpretation	87%	88%	Highest
Analysis	51%	76%	Moderate
Evaluation	49%	75%	Low
Inference	46%	72%	Lowest
Explanation	45%	76%	Low
Self-Regulation	50%	76%	Moderate

The data indicates a consistent pattern where students showed the highest ability in the interpretation indicator $(87 - 88\%)$ and the lowest in evaluation, inference, and explanation indicators $(45 - 51\%)$. This finding indicates that students are able to understand and explain the given information, but experience difficulties in assessing the correctness of solutions, drawing logical conclusions, and providing supporting reasons (Rosyidin, 2022). This condition reflects the limitations of Bloom's Taxonomy in classifying students' depth of understanding gradually, as Bloom focuses more on types of cognitive processes than on the structure of responses shown by students (Adeniji et al., 2022). In other words, students may be at the same cognitive level (for example C4 or C5) but have different qualities of understanding, which cannot be adequately revealed by Bloom's Taxonomy (Yuanita et al., 2023). Therefore, an assessment framework capable of distinguishing the quality of student responses more nuancedly is needed, one of which is SOLO Taxonomy (Svensäter & Rohlin, 2023).

The SOLO Taxonomy-based assessment framework proposed in this article offers a solution to these limitations by classifying student responses into five hierarchical levels that reflect depth of understanding gradually (Adeniji et al., 2022). At the prestructural level, students do not yet

have meaningful understanding and provide irrelevant responses, while at the unistructural level students are only able to use one piece of information or relevant aspect of a concept (Yuanita et al., 2023). The multistructural level reflects students' ability to mention several aspects separately, the relational level indicates the ability to integrate information into a coherent whole, and the extended abstract level is the highest level where students are able to generalize understanding and create innovative solutions (Evangelista & Rabut, 2025). With this framework, teachers can identify each student's position of understanding more accurately and provide specific feedback according to the level achieved (Ariani et al., 2026). Furthermore, the SOLO-based rubric adapting transition levels from Fernandez & Guzon (2025) enables more nuanced assessment of student responses that fall between two main levels.

The integration of waste management context into the SOLO-based assessment framework provides an additional dimension that enriches students' learning experiences and fosters environmental awareness (Giannini, 2020). This context allows students to apply the concepts of volume and surface area of cuboids in real situations, such as calculating waste bin capacity, determining material efficiency for making recycled containers, and designing optimal waste bin designs based on waste generation data (Sembiring et al., 2026; Asih et al., 2026). At the relational level, students can integrate volume and surface area concepts to solve contextual problems, while at the extended abstract level students are challenged to generalize understanding and create innovative solutions for waste management at school (Svensäter & Rohlin, 2023). This approach aligns with the principles of Education for Sustainable Development (ESD) which emphasizes the importance of connecting knowledge with real-world challenges and developing cross-sector competencies such as systems thinking, anticipatory thinking, and problem-solving (Olawale et al., 2025; Dao et al., 2026). Thus, assessment not only functions as a measuring tool but also as a means to build environmental awareness and students' social responsibility (Purnamasari et al., 2026).

The synthesis of these two research phases indicates that the journey from Bloom's Taxonomy to SOLO Taxonomy is not merely a change of theoretical framework, but an evolution in the way of viewing and measuring the quality of student understanding (Adeniji et al., 2022). Bloom's Taxonomy has provided a strong foundation in identifying the types of cognitive processes to be measured Toharudin et al. (2025), while SOLO Taxonomy complements it with the ability to classify the structure of student responses more nuancedly (Fernandez & Guzon, 2025). The integration of waste management context as a manifestation of ESD adds dimensions of meaningfulness and relevance, so that assessment not only measures cognitive abilities but also fosters environmental awareness and social responsibility (Cuesta et al., 2024; Stoett et al., 2024). Findings from the first research phase indicate that the Bloom-based instrument is valid, reliable, practical, and effective, but also reveals limitations in classifying students' depth of understanding (Priyanto et al., 2024). Therefore, the SOLO framework proposed in the second phase is expected to address these limitations and provide a more comprehensive picture of the quality of student understanding in geometry material (Yuanita et al., 2023; Andikos et al., 2025).

4. CONCLUSION

The synthesis of these two research phases concludes that the journey from Bloom's Taxonomy to SOLO Taxonomy reflects an evolution in measuring the quality of student understanding. The empirical findings from the first phase confirm that the developed Bloom-based summative assessment instrument is valid, reliable, practical, and effective, with an average validity score of 3.79, a reliability coefficient of 0.95, practicality scores of $86-89\%$, and

effectiveness scores ranging from 55% to 77%. However, the analysis also reveals a significant limitation: the instrument struggles to classify students' depth of understanding, particularly in the evaluation, inference, and explanation indicators, where achievement remains low (45 – 51%). This limitation underscores the need for a more nuanced assessment framework. Therefore, the formative and summative assessment framework based on SOLO Taxonomy with a waste management context proposed in the second phase is expected to address these limitations by integrating five hierarchical levels, a rubric with transition levels (Fernandez & Guzon, 2025) and an ESD context relevant to Phase D students' lives (Giannini, 2020; Purnamasari et al., 2026). The primary limitation of this article lies in the absence of empirical implementation of the proposed SOLO framework, meaning its validity and reliability have yet to be tested. Consequently, it is recommended for future researchers to implement this framework on a broader scale (Svensäter & Rohlin, 2023), test its effectiveness in improving students' critical thinking skills and environmental awareness (Ariani et al., 2026), and explore the integration of technology in SOLO-based assessment to enhance interactivity and accessibility (Andikos et al., 2025).

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the mathematics teachers and students at SMPN 1 Rambipuji who participated in the first phase of the research, as well as to the validators who provided valuable input for the development of the assessment instruments. Appreciation is also extended to the lecturers of the Mathematics Education Study Program for their guidance and support throughout the research process. This article is part of a larger research endeavor that bridges completed and planned studies, and the authors hope it contributes to the advancement of geometry assessment in Indonesia.

REFERENCES

Adeniji, S. M., Baker, P., & Schmude, M. (2022). Structure of the observed learning outcomes (solo) model: a mixed-method systematic review of research in mathematics education. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(6). https://doi.org/10.29333/ejmste/12087

Andikos, A. F., Rina, Sari, S. P., Amin, M., Mirdad, J., Afriwes, & Baroud, N. (2025). Development of solo based hots assessment in audio video learning for 3t vocational schools: Integrating religious moderation in mentawai. *IJORER : International Journal of Recent Educational Research*, *6*(4), 1081–1094. https://doi.org/10.46245/ijorer.v6i4.953

Ariani, F., Suyanto, Yuliana, L., Asriyah, & Rahmah, L. (2026). From taxonomies to rubrics: solo-based formative assessment to improve understanding of population dynamics. *International Journal of Learning, Teaching and Educational Research*, *25*(1), 518–535. https://doi.org/10.26803/ijlter.25.1.25

Asih, A. M. S., Prasetyo, D. H., & Sopha, B. M. (2026). Household recycling behavior and community based waste management: an agent based modeling of socioeconomic heterogeneity in urban indonesia. *Journal of Environmental Management*, *399*. https://doi.org/10.1016/j.jenvman.2026.128658

Asri, M. F., Cahyani, I., & Kurniawan, K. (2024). Utilization of indonesian e-modules based on the icare approach as an innovation in learning for students. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, *10*(3), 1181. https://doi.org/10.33394/jk.v10i3.12287

Audiwinanda, S., & Mahmudi, A. (2025). Validity and Reliability of Students' Mathematical Communication Instruments in Class VII Middle School Statistics Material. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *28*(1), 1–12. https://doi.org/10.21831/pep.v29i1.71234

Creswell, J. W. (2012). *Educational Research* (P. A. Smith, Ed.; Fourth). Pearson.

Cuesta, L., López-Fernández, C., Paños, E., & Ruiz-Gallardo, J.-R. (2024). Teachers' Attitudes Towards SDG-12 Responsible Consumption and Production. Development and Validation of a

Measurement Scale. *International Research in Geographical and Environmental Education*. https://doi.org/10.1080/10382046.2024.239949

Dao, H. N., Tho, L. N. A., & Tang, M. D. (2026). Developing Key Competencies for Sustainability Through Pre-Service Teachers' Design of SDG Related Mathematics Activities. *Journal on Mathematics Education*, *17*(1), 89–114. https://doi.org/10.22342/jme.v17i1.pp89-114

Evangelista, A. E., & Rabut, J. F. (2025). Enhancing Learning Outcomes: A Structure of Observed Learning Outcomes (SOLO) Taxonomy-Based Assessment Approach in Mathematics 7 in a Tri-People Dominated High School. *International Journal of Open-Access, Interdisciplinary and New Educational Discoveries of ETCOR Educational Research Center (IJOINED ETCOR)*, *4*(2). https://doi.org/10.63498/nxz2st268

Facione, P. (2015). *Critical Thinking: What It Is and Why It Counts*. https://www.researchgate.net/publication/251303244

Fadhli, R., Suharyadi, A., Firdaus, F. M., & Bustari, M. (2023). Developing a Digital Learning Environment Team Based Project to Support Online Learning in Indonesia. *International Journal of Evaluation and Research in Education*, *12*(3), 1599–1608. https://doi.org/10.11591/ijere.v12i3.24040

Fernandez, P. J. M., & Guzon, A. F. H. (2025). A SOLO Taxonomy-Based Rubric for Assessing Conceptual Understanding in Applied Calculus. *Journal on Mathematics Education*, *16*(2), 559–580. https://doi.org/10.22342/jme.v16i2.pp559-580

Giannini, S. (2020). *Education for Sustainable Development: A Roadmap*. UNESCO. https://doi.org/10.54675/YFRE14480

Hobri, Putri, M. A. K., Murtikusuma, R. P., Oktavianingtyas, E., & Pambudi, D. S. (2021). *Pengembangan lks Kelas 7 SMP aritmetika sosial dan penggunaanya dalam pembelajaran blended learning tipe flipped classroom*. *Edumatica Jurnal Pendidikan Matematika*, *11*(1). https://doi.org/https://doi.org/10.22437/edumatica.v11i01.12032

Kania, N., & Kyaruzi, F. (2026). Establishing quality indicators for higher order thinking skills: development and validation using factor analysis. *International Journal of Evaluation and Research in Education*, *15*(4), 3346–3361. https://doi.org/10.11591/ijere.v15i4.38884

Lafuente-Lechuga, M., Cifuentes-Faura, J., & Faura-Martínez, Ú. (2024). Teaching sustainability in higher education by integrating mathematical concepts. In *International Journal of Sustainability in Higher Education* (Vol. 25, Number 1, pp. 62–77). Emerald Publishing. https://doi.org/10.1108/IJSHE-07-2022-0221

Maruf, Marisda, D. H., Sultan, A. D., B, F. C. A., & L, W. D. (2024). Development of collaborative online learning model based on case method in optics courses to train creative and communication skills. *Jurnal Penelitian Pendidikan IPA*, *10*(9), 6655–6661. https://doi.org/10.29303/jppipa.v10i9.8448

Olawale, B. E., Hendricks, W., & Saddiq, K. (2025). Education for Sustainable Development and its Implication for the Preparation of Pre-Service Mathematics Teachers. In B. I. Omodan (Ed.), *Open Books and Proceedings* (6). University of the Free State. https://doi.org/10.38140/obp3-2025-06

Pankaew, K., Chanowan, S., & Piyakoson, S. (2025). Design and development of innovation towards contemporary educational innovation. *Asian Journal of Education and Training*, *11*(3), 88–95. https://doi.org/10.20448/edu.v11i3.6999

Priyanto, A., Hartoyo, A., Noviani, E., Matematika, P., & Tanjungpura, U. (2024). *Pengembangan asesmen pembelajaran matematika untuk menguji kemampuan bernalar kritis dan kreatif siswa Sekolah Menengah Atas*. *Journal of Education Research*, *5*(2), 1139–1146. https://doi.org/https://doi.org/10.37985/jer.v5i2.975

Purnamasari, R., Safitri, N., Handayani, R., & Pratama, M. A. (2026). Designing and validating a mathematics learning model integrating green economy and numeracy for ESD. *Discover Sustainability*, *7*(1). https://doi.org/10.1007/s43621-026-03056-5

Ramadhani, R., Nuraeni, E., & Purwianingsih, W. (2025). PISA Based Assessment Development in Environmental Issues: Fostering Numeracy Literacy Through Waste Management Context. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, *11*(3), 811–827. https://doi.org/10.22219/jpbi.v11i3.41082

Rosyidin, M. A. (2022). *Pengembangan instrumen asesmen berpikir tingkat tinggi pada matematika SMP materi bangun ruang sisi datar*. *Jurnal Pendidikan Matematika Dan Sains*, *10*(2), 127–134. https://doi.org/10.21831/jpms.v10i2.41129

Sembiring, E., Fenitra, R. M., Dangkua, A. R., Al Khoeriyah, Z. B., Van Der Laan, A. Z., Fan, Y., Ceschin, F., & Jobling, S. (2026). Exploring household waste management behavior through the stimulus organism response model. *Acta Psychologica*, *262*. https://doi.org/10.1016/j.actpsy.2025.106083

Stoett, P., Scrich, V. M., Elliff, C. I., Andrade, M. M., de M. Grilli, N., & Turra, A. (2024). Global Plastic Pollution, Sustainable Development, and Plastic Justice. In *World Development* (Vol. 184). Elsevier Ltd. https://doi.org/10.1016/j.worlddev.2024.106756

Svensäter, G., & Rohlin, M. (2023). Assessment Model Blending Formative and Summative Assessments Using the SOLO Taxonomy. *European Journal of Dental Education*, *27*(1), 149–157. https://doi.org/10.1111/eje.12787

Toharudin, T., & Dewi, L. (2025). *Panduan Mata Pelajaran Matematika Fase A-F dan Fase F Tingkat Lanjut*. Badan Standar, Kurikulum, dan Asesmen Pendidikan.

Toharudin, T., Dewi, L., Anggraena, Y., Ginanto, D. E., Kesuma, A. T., & Setiyowati, D. (2025). *Panduan Pembelajaran dan Asesmen (Revisi Ke-3). Badan Standar, Kurikulum, dan Asesmen Pendidikan (BSKAP) Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.*

Vásquez, C., Alsina, Á., Seckel, M. J., & García-Alonso, I. (2023). Integrating sustainability in mathematics education and statistics education: a systematic review. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(11). https://doi.org/10.29333/EJMSTE/13809

Yuanita, P., Maimunah, M., Kartini, K., & Arif, N. (2023). Analysis of solo taxonomy based test instruments in approaching the merdeka curriculum. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *8*(1), 12–29. https://doi.org/10.23917/jramathedu.v8i1.1270

Zana, F. M., Sa'dijah, C., & Susiswo. (2024). The cognitive alignment of mathematics teachers' assessments and its curriculum. *International Journal of Evaluation and Research in Education* , *13*(3), 1561–1575. https://doi.org/10.11591/ijere.v13i3.26814