About the Journal

Title: Edumatica: Jurnal Pendidikan Matematika

ISSN: P-ISSN 2088-2157 dan E-ISSN 2580-0779

Subject: Mathematics Education

Frequency Publish: Regularly three times a year in April, August, and December, Every issue consisted of 7 - 15 articles, and therefore, every volume consisted 21 - 45 articles.

Language: English

Indexed at: Edumatica has been indexed in SINTA 2 as of March 21, 2025, based on Decree No. 0173/C3/DT.05.00/2025, Crossref, DOAJ

Citation: Google Scholar

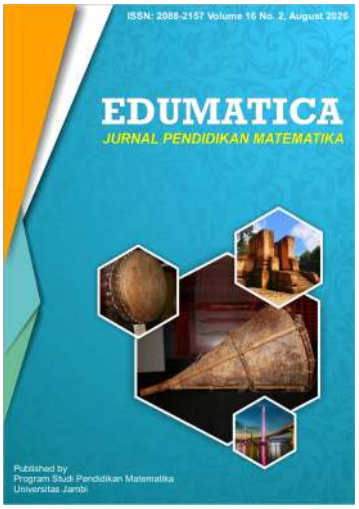
Archive preservation: Garuda, Pendidikan Matematika, FKIP Universitas Jambi

OAI: https://online-journal.unja.ac.id/edumatica/oai

Edumatica: Jurnal Pendidikan Matematika is a scholarly periodical. Edumatica publishes articles of research in mathematics education including teaching and learning, instruction, curriculum development, learning environments, teacher education, educational technology, educational developments, from many kinds of research such as survey, research and development, experimental research, classroom action research, etc. Edumatica first edition published in 2011. Edumatica: Jurnal Pendidikan Matematika are published regularly three times a year in April, August, and December. The journal maintained by Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Jambi, Indonesia. Edumatica: Jurnal Pendidikan Matematika cooperates with experienced writers in writing articles at the international level as reviewers and editorial teams. Authors who are interested in publishing their writings in Edumatica: Jurnal Pendidikan Matematika can do Submission of articles by logging in first. While writers who do not have an account can register first and make sure the email is used actively.

CURRENT ISSUE

Vol. 16 No. 2 (2026): Edumatica: Jurnal Pendidikan Matematika (August 2026)



This issue has been available online since **August 31, 2026**. All articles in this issue (... **original research articles**) were authored/co-authored from **... countries** (..., ..., and ...)**,** ... **provinces** (S..., ..., and ...), and ... **institutions (..., ..., and ...).**

PUBLISHED: 2026-08-31

Articles

PDF

Development of an Interactive Learning Medium to Improve Students' Abstraction Ability

Swasti Maharani, Hendrisa Rizqie Romandoni, Ardiana Ardiana

37-52

PDF

The Developmental Mathematical Literacy Instruments for Junior High School Teachers: Construction and Validation Study

Niluh Sulistyani, Albertus Hariwangsa Panuluh, Kintan Limianingsih, Margaretha Madha Melissa

53-64

PDF

How AI Facilitates Students in Constructing Ideas To Solve Mathematical Problems: A Critical Thinking Perspective

Dessy Wulandari, Ali Shodikin

Enhancing Numeracy Skills through a GeoGebra-Integrated Collaborative Platform in Online Mathematics Learning

Dewi Iriani, Khairul Anwar, Sofnidar Sofnidar

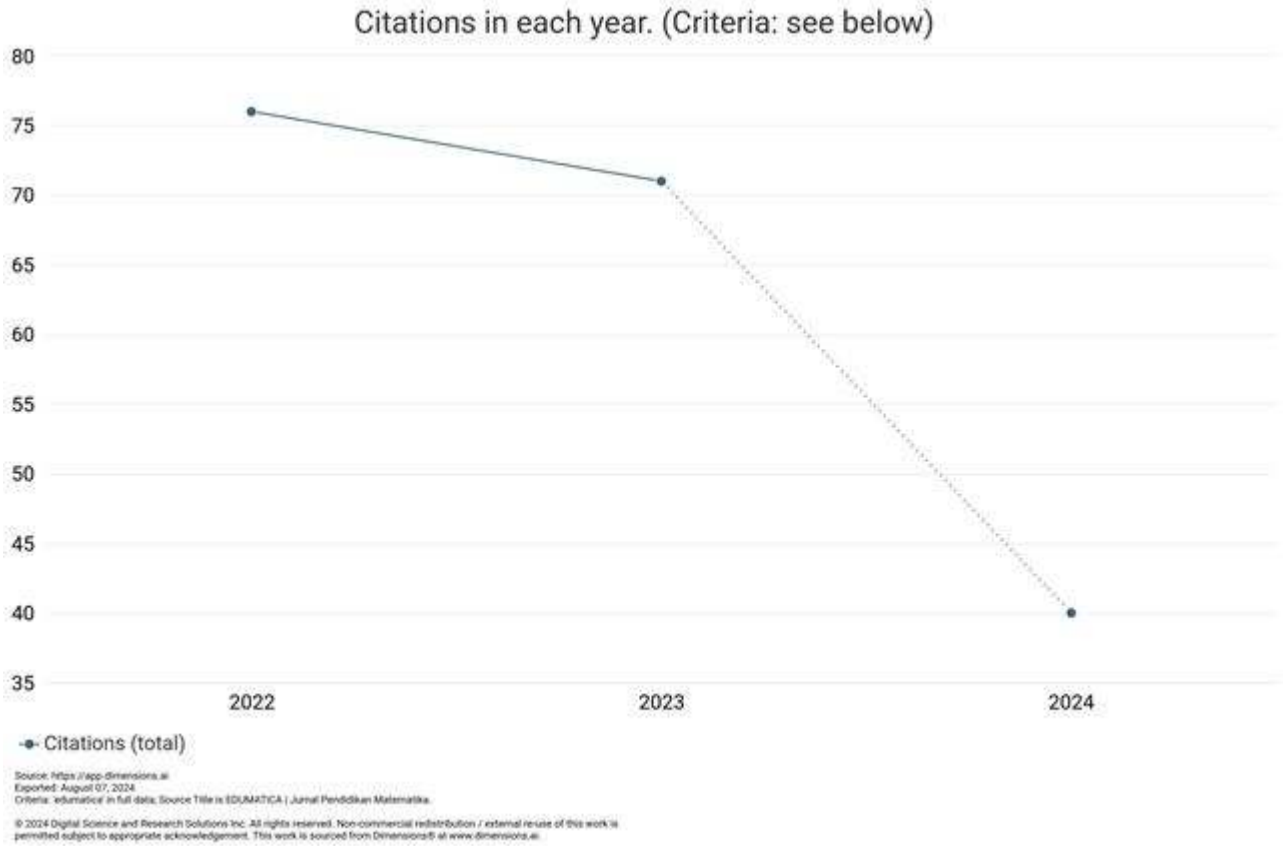
Preservice Mathematics Teachers' Contextual Numeracy: Difficulty and Changes in Problem-Solving Pattern

Sri Winarni, Kamid Kamid, Asrial Asrial, Jefri Marzal

The Ability to Prove of Real Analysis Problems: Investigation with APOS Conception

Muhammad Win Afgani, Retni Paradesa, Yuli Fitrianti

VIEW ALL ISSUES ›



MAKE A SUBMISSION

ABOUT JOURNAL

FOCUS AND SCOPE

AUTHOR GUIDELINE

Kementerian Pendidikan Tinggi, Sains, dan Teknologi

SERTIFIKAT

Akreditasi Jurnal

TERAKREDITASI

No. SK : 10/C/C3/DT.05.00/2025 Tanggal : 21 Maret 2025

Direktur Penelitian, dan Pengabdian kepada Masyarakat

dengan ini memberikan kepada

Edumatica: Jurnal Pendidikan Matematika

25800779

EISSN : 25800779

Publisher : Universitas Jambi

Peringkat Akreditasi Jurnal Ilmiah Periode I 2025

Akreditasi Ulang di Peringkat 2 mulai

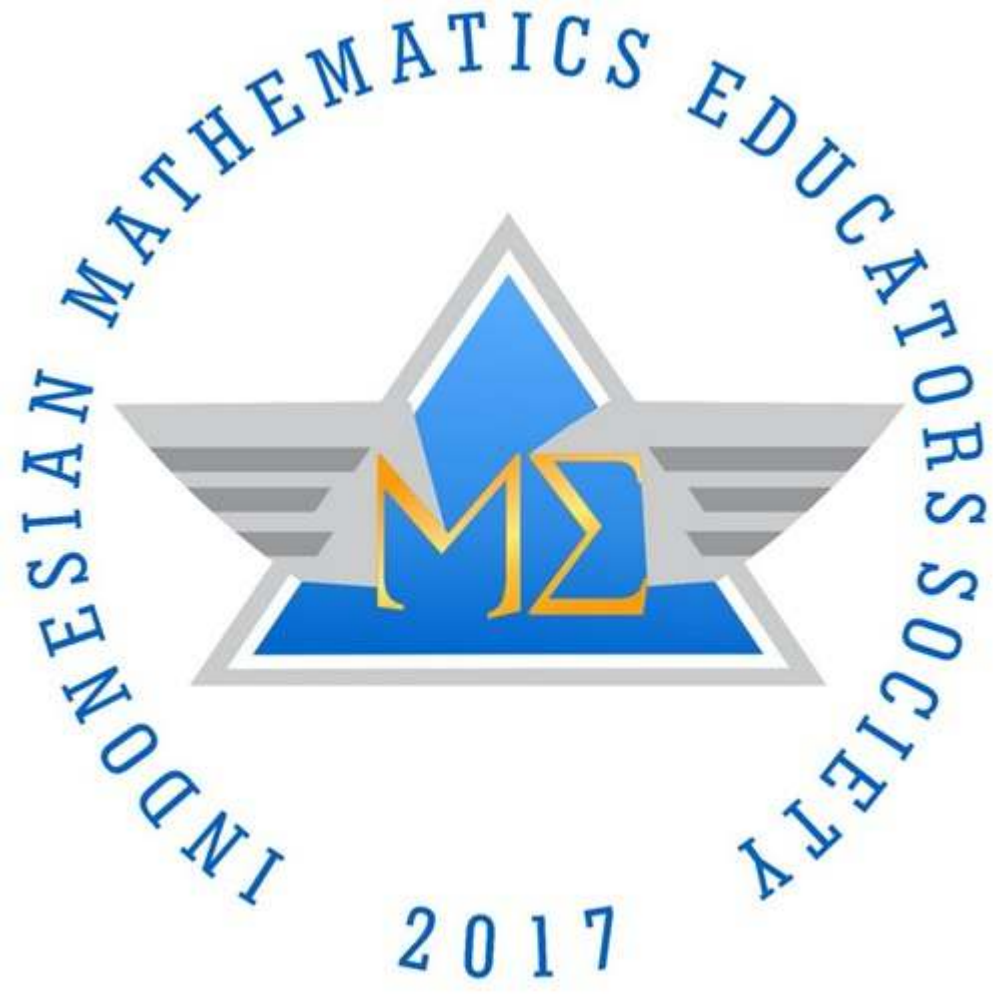
Volume 14 Nomor 1 Tahun 2024 sampai Volume 18 Nomor 2 Tahun 2028

I Ketut Adnyana

NIP 196805151994031004

Article Template Edumatica

COLLABORATE WITH



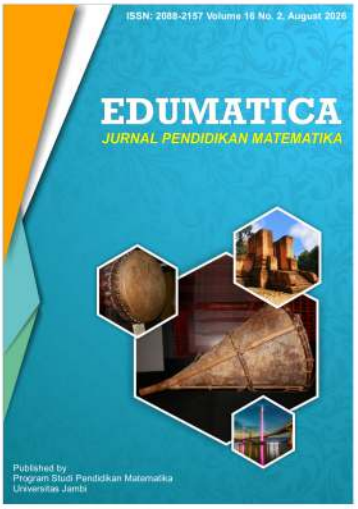
PKS Download Here

PKS Download Here

flag

EDUMATICA
Jurnal Pendidikan Matematika
P-ISSN 2088-2157 E-ISSN 2580-0779

Vol. 16 No. 2 (2026): Edumatica: Jurnal Pendidikan Matematika (August 2026)



This issue has been available online since **August 31, 2026**. All articles in this issue (... **original research articles**) were authored/co-authored from **... countries** (..., ..., and ...)**,** ... **provinces** (S..., ..., and ...), and ... **institutions (..., ..., and ...).**

PUBLISHED: 2026-08-31

ARTICLES

Pattern

Sri Winarni, Kamid Kamid, Asrial Asrial, Jefri Marzal

The Ability to Prove of Real Analysis Problems: Investigation with APOS Conception

Muhammad Win Afgani, Retni Paradesa, Yuli Fitrianti

MAKE A SUBMISSION

ABOUT JOURNAL

FOCUS AND SCOPE

AUTHOR GUIDELINE

REVIEWER

SECTION POLICIE

PUBLICATION FREQUENCY

REVIEW POLICY

RIGHTS AND LICENCES

OPEN ACCES STATEMENT

ARCHIVING

PLAGIARISM AND RETRACTION

ARTICLES PROCESSING CHARGE

Kementerian Pendidikan Tinggi, Sains, dan Teknologi

SERTIFIKAT

Akreditasi Jurnal

TERAKREDITASI

No. SK : 10/C/C3/DT.05.00/2025 Tanggal : 21 Maret 2025

Direktur Penelitian, dan Pengabdian kepada Masyarakat

dengan ini memberikan kepada

Edumatica: Jurnal Pendidikan Matematika

25800779

EISSN : 25800779
Publisher : Universitas Jambi

Peringkat Akreditasi Jurnal Ilmiah Periode I 2025

Akreditasi Ulang di Peringkat 2 mulai
Volume 14 Nomor 1 Tahun 2024 sampai Volume 18 Nomor 2 Tahun 2028

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
DIREKTORAT PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT

I Ketut Adnyana
NIP 196805151994031004

COLLABORATE WITH

PKS Download Here

PKS Download Here

flag

EDUMATICA
Jurnal Pendidikan Matematika
P-ISSN 2088-2157 E-ISSN 2580-0779

The **Developmental Mathematical Literacy Instruments for Junior High School Teachers:**

Construction and Validation Study

Niluh Sulistyani

Sanata Dharma University

https://orcid.org/0000-0001-5506-7198 (unauthenticated)

Albertus Hariwangsa Panuluh

Sanata Dharma University

https://orcid.org/0000-0003-1393-5713 (unauthenticated)

Kintan Limianingsih

Sanata Dharma University

https://orcid.org/0009-0004-6608-4152 (unauthenticated)

Margaretha Madha Melissa

International Education Development Program, Hiroshima University

https://orcid.org/0009-0003-4260-4005 (unauthenticated)

DOI: https://doi.org/10.22437/edumatica.v16i2.53077

Keywords: Design-Based Research, Junior high school, Mathematical literacy, Teacher, Tessmer

ABSTRACT

Teachers are required to facilitate students' mathematical literacy, but teachers' literacy skills are rarely considered due to the lack of instruments for assessing teachers' mathematical literacy. This study aims to develop a mathematical literacy instrument for teachers, focusing on the processes of instrument

construction and validation. This Design-Based Research (DBR) adopts the Tessmer formative evaluation development model, consisting of preliminary, self-evaluation, expert reviews, one-to-one, small-group, and field tests. This study involved two experts as instrument validators (logical validity) and a total of 60 junior high school teachers in Yogyakarta Province to assess readability at the one-to-one stage and in small groups as subjects of field tests (empirical validation). The logical validation results were analyzed using the Aiken V index; readability test results were analyzed descriptively; and the results of empirical validation were analyzed using SPSS. The instrument construction process follows the principles of educational assessment development within the PISA framework and produces 25 literacy items, each involving 10 stimuli with varying contexts. The average V Aikens of 0.95 indicates a valid instrument, while the average readability test score of 4.14 indicates the instrument is in the good readability category. Based on the field test results, each item is valid because it has a Sig value ≤ 0.05. Overall, the question instrument is reliable, with a Cronbach's alpha (α) of 0.860 > 0.7. The instrument can measure mathematical literacy skills among teachers and assess all levels of thinking and question difficulty.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131–142. https://doi.org/https://doi.org/10.1177/0013164485451012

Aisyah Aisyah, D. J. (2022). The description of Indonesian students' mathematics literacy in the last decade. International Journal of Trends in Mathematics Education Research. https://doi.org/https://doi.org/10.33122/ijtmer.v5i1.114.

Alfaruqi, A. Z., & Nurwahidayah. (2025). Jurnal Pendidikan IPS. 15(1), 11–19. https://doi.org/https://doi.org/10.37630/jpi.v15i1.2559

Amin, B., Baharuddin, B., Nur, F., Nursalam, N., & Angriani, A. D. (2025). Pengembangan Tes Asesmen Kompetensi Mahasiswa Untuk Mengukur Kemampuan Literasi Numerasi. AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 14(1), 54. https://doi.org/10.24127/ajpm.v14i1.9494

Apriandi;, D., Murtafiah;, W., Suprapto;, E., Masfingatin;, T., & Lusiana, R. (2024). The development of problem-based calculus E-modules to improve students' numeracy literacy skills. AIP Conference Proceedings, 3148(1). https://doi.org/https://doi.org/10.1063/5.0243257

Benson, J., & Clark, F. (1982). A Guide for Instrument Development and Validation. The American Journal of Occupational Therapy, 36(12), 789–800. https://doi.org/https://doi.org/10.5014/ajot.36.12.789

Blum, W., & Leiss, D. (2020a). The Role of Contexts in Mathematical Modeling – Towards a Conceptual Framework. ZDM Mathematics Education, 52, 149–163. https://doi.org/10.1007/s11858-019-01100-1

Chen, X., Zhou, J., Wang, J., Wang, D., Liu, J., Shi, D., Yang, D., & Pan, Q. (2022). Visualizing Status, Hotspots, and Future Trends in Mathematical Literacy Research via Knowledge Graph. Sustainability (Switzerland), 14(21). https://doi.org/10.3390/su142113842

Chiu, T. K. F., Chen, Y., Yau, K. W., Chai, C. Sing, Meng, H., King, I., Wong, S., & Yam, Y. (2024). Developing and validating measures for AI literacy tests: From self-reported to objective measures. Computers and Education: Artificial Intelligence, 7(August), 100282. https://doi.org/10.1016/j.caeai.2024.100282

Citrawan, I. W., et al. (2024). Special Education Teachers ' Ability in Literacy and Numeracy Assessments Based on Local Wisdom. 8(1), 145–157. https://doi.org/https://doi.org/10.22437/jiituj.v8i1.32608

DeVellis, R. F. (2003). Scale Development: Theory and Applications (Second Edition). SAGE Publications.

Golla, E., & Reyes, V. (2022b). Developing Mathematical Literacy Test Items for Junior High School Teachers Based on the PISA Framework. International Journal of Evaluation and Research in Education, 11(4), 2045–2058. https://doi.org/10.11591/ijere.v11i4.21983

Helstad, K., Solbrekke, T. D., & Wittek, A. L. (2017). Exploring teaching academic literacy in mathematics in teacher education. Education Inquiry, 8(4), 318–336. https://doi.org/10.1080/20004508.2017.1389225

Hilmi, I. &

Kismiantini. (2024). Teacher Competencies, The Shortage of School Resources, and Mathematics Achievement Based on PISA 2018 Indonesia. Jurnal Pendidikan Matematika, 18(2). https://doi.org/https://doi.org/10.22342/jpm.v18i2.pp245-258

Holenstein, M., Bruckmaier, G., & Grob, A. (2021). Transfer effects of mathematical literacy: an integrative longitudinal study. European Journal of Psychology of Education, 36(3), 799–825. https://doi.org/10.1007/s10212-020-00491-4

Hwang, J., & Ham, Y. (2021). Relationship between mathematical literacy and opportunity to learn with different types of mathematical tasks. Journal on Mathematics Education, 12(2), 199–222. https://doi.org/10.22342/JME.12.2.13625.199-222

Kandil, S., & Işıksal-Bostan, M. (2019). Effect of inquiry-based instruction enriched with origami activities on achievement and self-efficacy in geometry. International Journal of Mathematical Education in Science

and Technology, 50(4), 557–576. https://doi.org/10.1080/0020739X.2018.1527407

Kappassova, S., Abylkassymova, A., Bulut, U., Zykrina, S., Zhumagulova, Z., & Balta, N. (2025). Mathematical literacy and its influencing factors: A decade of research findings (2015-2024). Eurasia Journal of Mathematics, Science and Technology Education, 21(7). https://doi.org/10.29333/ejmste/16615

Kioupi, V., & Voulvoulis, N. (2019). Education for sustainable development: A systemic framework for connecting the SDGs to educational outcomes. Sustainability (Switzerland), 11(21). https://doi.org/10.3390/su11216104

Manfreda Kolar, V.M, Hodnik, T. (2021). Mathematical Literacy from the Perspective of Solving Contextual Problems. European Journal of Educational Research, 10(1), 467–483. https://doi.org/10.12973/eu-jer.10.1.467

McKay, V. (2018). Literacy, lifelong learning, and sustainable development. Australian Journal of Adult Learning, 58(3), 390–425. https://ala.asn.au/wp-content/uploads/2021/01/AJAL-58-3-Full.pdf

McKenney, S., & Reeves, T. C. (2018). About educational design research. Conducting Educational Design Research, 5–33. https://doi.org/10.4324/9781315105642-3

Merino-Soto, C. (2023). Coeficientes V de Aiken: diferencias en los juicios de validez de contenido. MHSalud: Revista en Ciencias del Movimiento Humano y Salud, 20(1), 1–10. https://doi.org/10.15359/mhs.20-1.3

Miles, M. B.; Huberman, M. S. J. (2014). Qualitative Data Analysis (3rd ed.). SAGE Publication.

Mislevy, R. J., Haertel, G., & Almond, R. G. (2021). Evidence-Centered Design for Educational Assessment: A Research Update. Educational Measurement: Issues and Practice, 40(1), 16–31. https://doi.org/10.1111/emip.12424

Muhaimin, L. H., Sholikhakh, R. A., Yulianti, S., Ardani, H., A., & Sahara, S. (2024). Unlocking the secrets of students' mathematical literacy to solve mathematical problems: A systematic literature review. Eurasia Journal of Mathematics, Science and Technology Education, 20(4). https://doi.org/10.29333/ejmste/14404

Nadila, S. A., & Lestiana, H. T. (2024). Kemampuan Literasi Matematika Siswa Berdasarkan Gaya Belajar. Edumatica: Jurnal Pendidikan Matematika, 14(03). https://doi.org/https://doi.org/10.22437/edumatica.v14i3.38303

Nieveen, N., & Folmer, E. (2013). Formative Evaluation in Educational Design Research. In Educational Design Research – Part A: An Introduction (pp. 152–169). Netherlands Institute for Curriculum

Development (SLO).

Nitko, A. J., & Brookhart, S. M. (2022). Educational Assessment of Students (8th ed.). Pearson Education.

OECD. (2016). Program for International Student Assessment (PISA) Results From PISA 2015. https://www.oecd.org/pisa/PISA-2015-Indonesia.pdf

OECD. (2019). PISA 2018 Results. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf

OECD. (2021). PISA 2021 Mathematics Framework. https://www.oecd.org/pisa/pisaproducts/pisa-2021-mathematics-framework.pdf

OECD. (2023). PISA 2022 Assessment and Analytical Framework. In Paris: Journal of OECH Publishing.

PISA. (2023). PISA 2022 Results Factsheets Indonesia. The Language of Science Education, 1, 1–9. https://oecdch.art/a40de1dbaf/C108.

Pulsiri, N., Vatananan-Thesenvitz, R., Tantipisitkul, K., Aung, T. H., Schaller, A. A., Schaller, A. M., Methananthakul, K., & Shannon, R. (2019). Achieving sustainable development goals for people with disabilities through digital technologies. PICMET 2019 - Portland International Conference on Management of Engineering and Technology: Technology Management in the World of Intelligent Systems, Proceedings, (November 2020). https://doi.org/10.23919/PICMET.2019.8893725

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

Sinclair, N., Bartolini Bussi, M. G., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: an ICME-13 survey team report. ZDM, 48(5), 691–719. https://doi.org/10.1007/s11858-016-0796-6

Stacey, K. (2011). The PISA view of mathematical literacy in Indonesia. Journal on Mathematics Education, 2(2), 95–126. https://doi.org/10.22342/jme.2.2.746.95-126

Stacey, K. & Turner, R. (2015). Assessing Mathematical Literacy: The PISA Experience (R. Stacey, Kaye & Turner, Ed.). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-319-10121-7

Suciati, S. Munadi, Sugiman, & Wiwin Dwi Ratna Febriyanti. (2020). Design and Validation of Mathematical Literacy Instruments for Assessment for Learning in Indonesia. European Journal of Educational Research, 9(2), 865–875. https://doi.org/https://doi.org/10.12973/eu-jer.9.2.865

Sugiyono. (2019). Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D. Alfabeta.

Tessmer, M. (2005). Planning and Conducting Formative Evaluations: Improving the Quality of Education and Training. Routledge.

Tout, D., Gal, I., & Schlöglmann, W. (2023). Mathematical Reasoning and Literacy: Revisiting PISA's Conceptualization for the Next Decade. Studies in Educational Evaluation, 78, 102210. https://doi.org/10.1016/j.stueduc.2023.102210

Tout, D., Coben, D., Geiger, V., Ginsburg, L., Hoogland, K., Maguire, T., Thomson, S., & Turner, Ross. (2017). Review of the PIAAC Numeracy Assessment Framework: Final Report.

Treagust, D. F., & Chandrasegaran, A. L. (2020). Developing Diagnostic Assessment Items to Evaluate Students' Understanding of Science and Mathematics Concepts. International Journal of Science and Mathematics Education, 18, 689–708. https://doi.org/10.1007/s10763-019-09971-0

Umbara, U., & Suryadi, D. (2019). Re-interpretation of mathematical literacy based on the teacher's perspective. International Journal of Instruction, 12(4), 789–806. https://doi.org/10.29333/iji.2019.12450a

UNESCO. (2024). Global report on teachers: addressing teacher shortages and transforming the profession. In the Global report on teachers: addressing teacher shortages and transforming the profession. https://doi.org/10.54675/figu8035

Wijaya, A., Retnawati, H., & Nusantara, T. (2024). Enhancing Teachers' Competence in Developing Mathematical Literacy Assessment Aligned with the PISA Framework. Eurasia Journal of Mathematics, Science and Technology Education, 20(2), em2295. https://doi.org/10.29333/ejmste/13670

Wijaya, A., Yogyakarta, U. N., Dewayani, S., Litara, Y., Effendi, A., Yogyakarta, U. N., & Gunawan, H. (n.d.). Framework Asesmen Kompetensi Minimum (AKM).

Wong, V., Yun, S., Ulang, N., & Husain, S. H. (2023). Measuring the Internal Consistency and Reliability of the Hierarchy of Controls in Preventing Infectious Diseases on Construction Sites : The Kuder-Richardson (KR-20) and Cronbach's Alpha. 1(1), 392–405. https://doi.org/https://doi.org/10.37934/araset.33.1.392405

Yusmar, F., & Fadilah, R. E. (2023). Analisis Rendahnya Literasi Sains Peserta Didik Indonesia: Hasil Pisa Dan Faktor Penyebab. LENSA (Lentera Sains): Jurnal Pendidikan IPA, 13(1), 11–19. https://doi.org/10.24929/lensa.v13i1.283

ISSN: 2088-2157 Volume 16 No. 2, August 2026

EDUMATICA

JURNAL PENDIDIKAN MATEMATIKA

Published by
Program Studi Pendidikan Matematika
Universitas Jambi

PDF

PUBLISHED

2026-10-03

ISSUE

Vol. 16 No. 2 (2026): Edumatica: Jurnal Pendidikan Matematika (August 2026)

SECTION

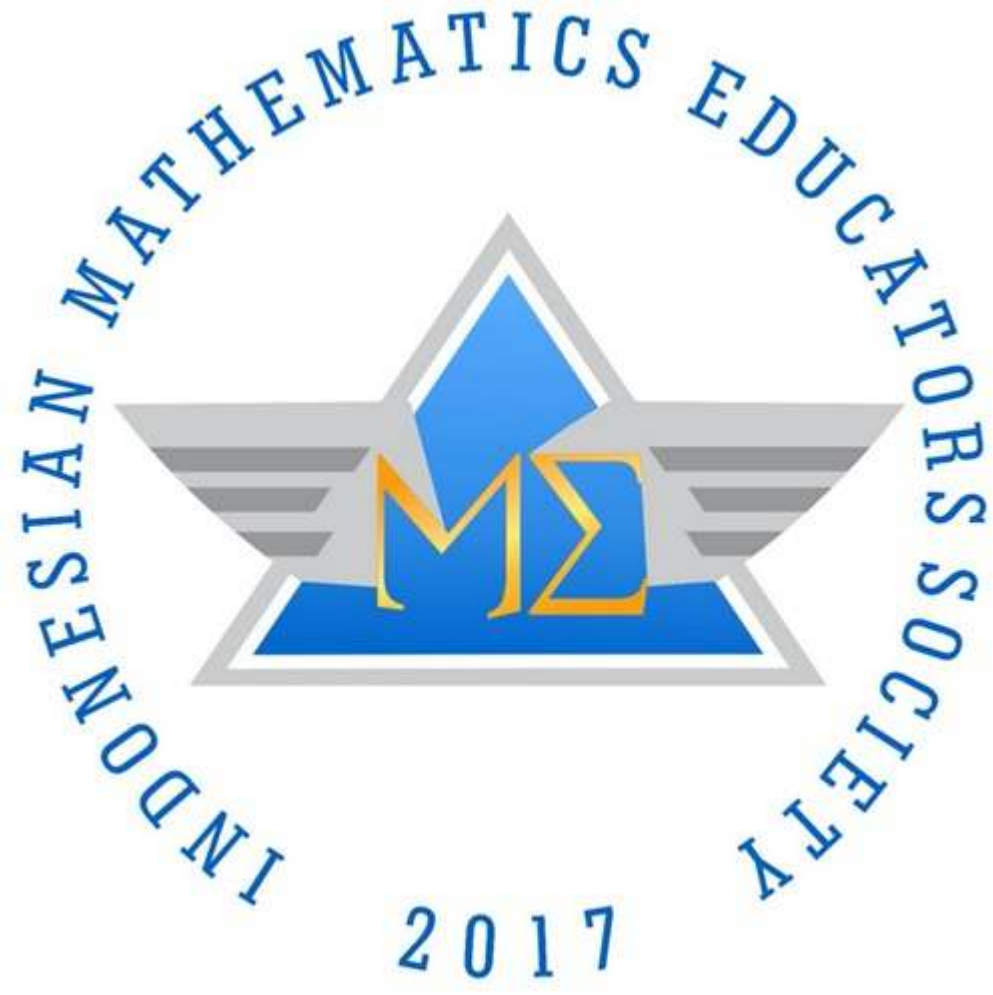
Articles

HOW TO CITE

Sulistyani, N., Panuluh, A. H., Limianingsih, K., & Melissa, M. M. (2026). The Developmental Mathematical Literacy Instruments for Junior High School Teachers: : Construction and Validation Study. *Edumatica : Jurnal Pendidikan Matematika*, *16*(2), 53-64. https://doi.org/10.22437/edumatica.v16i2.53077

Article Template Edumatica

COLLABORATE WITH



PKS Download Here

PKS Download Here

flag

Edumatica: Jurnal Pendidikan Matematika
Volume 16 Number 2 Agustus 2026

Submited : 01 January 2026
Revised : 11 February 2026
Accepted : 1 September 2026

DOI: https://doi.org/10.22437/edumatica.vl6i2.53077

The Developmental Mathematical Literacy Instruments for Junior High School Teachers: Construction and Validation Study

Niluh Sulistyani[1*], Albertus Hariwangsa Panuluh[2], Kintan Limianingsih[3], Margaretha Madha Melissa[4]
[1,2,3,4] Universitas Sanata Dharma Yogyakarta, Indonesia
[3] Universitas Sebelas Maret Surakarta, Indonesia
[4] International Education Development Program, Hiroshima University, Japan
E-mail*: niluh@usd.ac.id

Abstract

Teachers are required to facilitate students' mathematical literacy, but teachers' literacy skills are rarely considered due to the lack of instruments for assessing teachers' mathematical literacy. This study aims to develop a mathematical literacy instrument for teachers, focusing on the processes of instrument construction and validation. This Design-Based Research (DBR) adopts the Tessmer formative evaluation development model, consisting of preliminary, self-evaluation, expert reviews, one-to-one, small-group, and field tests. This study involved two experts as instrument validators (logical validity) and a total of 60 junior high school teachers in Yogyakarta Province to assess readability at the one-to-one stage and in small groups as subjects of field tests (empirical validation). The logical validation results were analyzed using the Aiken V index; readability test results were analyzed descriptively; and the results of empirical validation were analyzed using SPSS. The instrument construction process follows the principles of educational assessment development within the PISA framework and produces 25 literacy items, each involving 10 stimuli with varying contexts. The average V Aikens of 0.95 indicates a valid instrument, while the average readability test score of 4.14 indicates the instrument is in the good readability category. Based on the field test results, each item is valid because it has a Sig value ≤ 0.05. Overall, the question instrument is reliable, with a Cronbach's alpha (α) of $0.860 > 0.7$. The instrument can measure mathematical literacy skills among teachers and assess all levels of thinking and question difficulty.

Key Words: Desain-based research (DBR); Junior high school; Mathematical literacy; Teacher; Tessmer

INTRODUCTION

Mathematical literacy has become a global focus of attention because it is a crucial skill needed in everyday life (Holenstein et al., 2021) in the 21st century. Mathematical literacy is not just the ability to calculate (Kappassova et al., 2025), but also about the ability to reason (OECD, 2016, 2019, 2023), solve non-routine problems, and apply mathematical knowledge in a variety of everyday contexts (OECD, 2016; Tout et al., 2023); Omiralievna, 2024). These skills are essential for modern society to solve problems that require critical reasoning (Manfreda Kolar, V.M; Hodnik, 2021) as well as to read and make data-driven decisions on social, economic, and environmental issues (Chen et al., 2022), for example, climate change (Pulsiri et al., 2019). Furthermore, mathematical literacy is a key pillar in realizing the Sustainable Development Goals (SDGs) (McKay, 2018), which, in turn, contributes to holistic sustainable development (Kioupi & Voulvoulis, 2019). The existence of PISA, the Program for International Assessment, which measures students' mathematical literacy skills in countries belonging to the Organization for Economic Co-operation and Development (OECD), further underscores that mathematical literacy is an international concern (OECD, 2016).

Although essential, Indonesian students' mathematical literacy needs improvement. PISA results show that Indonesian students' mathematical literacy scores are far below the OECD average and have declined over the past three rounds, particularly in mathematics literacy. However, Indonesia's ranking in 2022 is better than in the previous period (OECD, 2023). Research findings at the elementary through high school levels also indicate that Indonesian students' mathematical literacy remains low (Aisyah Aisyah, 2022). A review of research in Indonesia also concluded that students' mathematical literacy skills remain low (Muhaimin et al., 2024).

The achievement of mathematical literacy among Indonesian students, as indicated by PISA results, is influenced by various factors (Yusmar & Fadilah, 2023), one of which is teacher ability and performance (Hwang & Ham, 2021). Teacher competence (pedagogical, moral, and instructional support) is significantly related to students' PISA mathematics scores (Hilmi & Kismiantini, 2024). The literacy skills of teachers in various regions of Indonesia vary widely, as do the challenges and obstacles they face (Citrawan et al., 2024). There is a significant relationship between low teacher competence and student academic achievement, including student literacy skills (Alfaruqi & Nurwahidayah, 2025). Teachers are required to foster mathematical literacy in students, but paradoxically, they themselves often lack adequate instruments to measure and develop these skills.

Research to develop mathematical literacy instruments in Indonesia has been extensive, but most of it focuses on students. Nadila & Lestiana (2024) designed PISA-based items that reflect Indonesian culture for junior high school students. Chiu et al (2024) developed an AI-supported instrument for students. Apriandi et al (2024) and Amin (2025) focused on preservice teachers through instruments related to calculus and AKM topics. Furthermore, existing teachers' mathematical literacy instruments are limited to measuring pedagogical skills, as UNESCO has noted in relation to teachers' challenges in mastering mathematical and scientific literacy (UNESCO, 2024), or to measuring their understanding of routine mathematical content (Kandil & Işıksal-Bostan, 2019). For example, research by Helstad on teachers' pedagogical skills indicates a tendency to ask more procedural questions, using routine questions that measure numeracy, as in traditional mathematics instruction. However, such instruction poses greater challenges for developing students' mathematical literacy skills (Helstad et al., 2017). While previous research focused on instruments measuring teachers' mathematical literacy insights (Umbara & Suryadi, 2019), this study specifically offers an instrument to measure teachers' mathematical literacy skills. This instrument is designed to measure teachers' ability to solve non-routine problems or problems requiring a high level of mathematical literacy that are rarely used.

This research offers a novel contribution by developing a mathematical literacy instrument and conducting a comprehensive construction and content validation process specifically designed for junior high school mathematics teachers. The results of this mathematical literacy instrument development will not only serve as a measurement tool but also as a diagnostic tool to identify areas where teachers need professional development. In addition, this instrument is expected to serve as a model for teachers to create similar questions in class.

METHOD

This study employs a Design-Based Research (DBR) approach, utilizing Tessmer's Formative Evaluation Model. Tessmer's model includes preliminary study, self-evaluation, expert review, one-to-one, small group, and field testing (Tessmer, 2005).

Participants included 60 junior high school mathematics teachers from the Yogyakarta Special Region (DIY), Indonesia, and two mathematics education lecturers who acted as expert validators. The first validator is a doctor of mathematics education with a Ph.D. and research experience related to PISA. The second validator is a lecturer and author of a book on numeracy literacy. The depth, track record, and expertise of the assessors are more crucial than the number of assessors (Diamond, 2014; Shrotryia & Dhanda, 2019). Following the small-group phase, the research continued with a forum group discussion involving two math teachers to further confirm the small-group results. The research procedure, adapted from Tessmer's model, is summarized in Figure 1

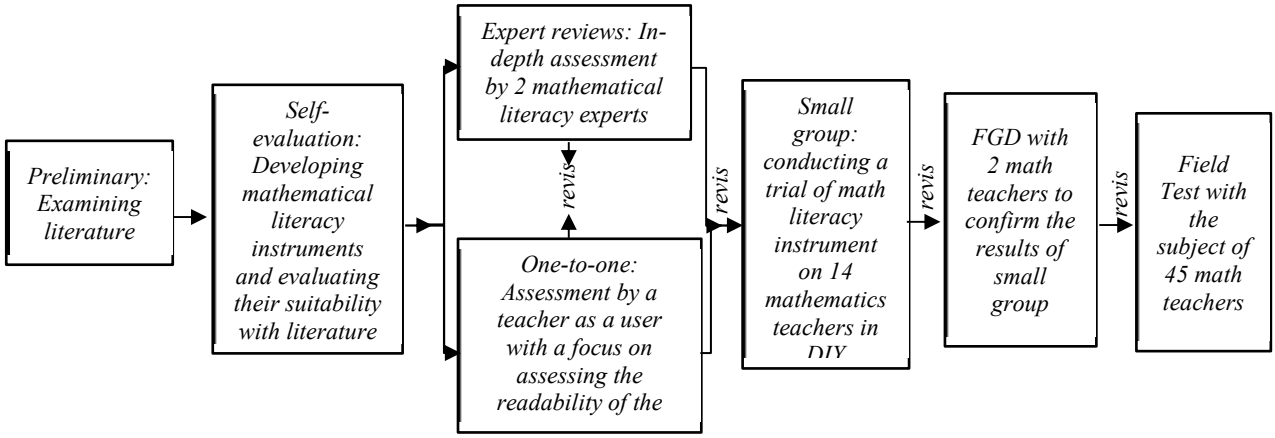


Figure 1. Research Stages: Adopting Tessmer's Model

In a preliminary stage, researchers reviewed the PISA mathematical literacy, curriculum analysis, and teacher needs assessment to ensure contextual relevance (Nieveen, N., & Folmer, 2013). In the self-evaluation stage, researchers created a draft instrument that began with the development of the context and test items, and then evaluated their suitability within the PISA framework (McKenney & Reeves, 2018). Two experts validated the draft instrument for an assessment using a validation sheet covering content, construct, language, and norms (McKenney & Reeves, 2018; Tessmer, 2005). Concurrently with the validation process, the draft questions were assessed by a mathematics teacher to evaluate the readability of the instrument's context and questions (Nieveen, N., & Folmer, 2013) using the instrument readability sheet. Input from the validator and readability experts served as the basis for revising the instrument. The revised instrument was subjected to a small-group test conducted by 14 teachers to determine the difficulty index (Tessmer, 2005) and to assess the practicality of the questions using a practicality questionnaire. The results were confirmed through FGDs with 2 teachers from the small-group subjects. The instrument was then administered to 45 teachers to empirically assess the quality of the questions as part of the field test.

Items on the validation sheet, readability sheet, and practicality questionnaire were rated on a 5-point Likert scale. The results of the content validity test from the two validators were analyzed using Aiken's V-index (Merino-Soto, 2023) with the formula:

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

where $s = r - l_0$ (r is the validator's score, l_0 is the lowest score on the scale), c is the number of rating categories (c = 5), and n is the number of validators (n = 2). The denominator in formula (1) is 8. Based on Aiken's (Aiken, 1985) significance table, the absolute minimum value that must be met to achieve statistical significance at α = 0.05 is 0.88.

The readability of the one-to-one stage and the practicality of the small-group stage were analyzed by calculating the average score, and the averages were then categorized according to Table 1—Readability and Practicality Criteria.

Table 1. *Readability and Practicality Criteria*

Mean Score Range	Category	Recommendation
4.21–5.00	Excellent	No revision needed
3.41–4.20	Good	Minor revision needed
2.61–3.40	Fair	Moderate revision needed
1.81–2.60	Poor	Major revision needed
1.00–1.80	Very Poor	Complete overhaul needed

(Sugiyono, 2019)

A small-group test with 14 teachers was conducted to assess item difficulty and to support the descriptive analysis. Teachers completed questionnaires and gave written feedback to refine the instrument. Their open-ended responses were analyzed qualitatively following Miles and Huberman's three-step process: data reduction, data display, and conclusion drawing, along with insights from focus group discussions (2014).

The questionnaire was administered to 45 teachers as part of the field test. Validity and reliability tests were conducted at this stage as part of the questionnaire's development phase, as described by Benson & Clark (1982). The validity analysis of the questions used the Pearson Product-Moment Correlation (questions were valid if Sig ≤ 0.05), and reliability was assessed using Cronbach's alpha. An α value > 0.70 was considered acceptable (DeVellis, 2003). Both were analyzed using SPSS software.

RESULTS

Preliminary

The instrument development process began with a review of the mathematical literacy literature, following the PISA framework, as a preliminary stage. This process took into account the literacy assessment development principles of Stacey, Kaye, and Turner (2015), which begin with analyzing mathematical literacy constructs, identifying real-world contexts, determining competencies, designing a blueprint, and writing draft test items. The preliminary process and results are shown in Table 2.

Table 2. *Mathematical Literacy Instrument Construction Process*

Step	Study Results	Basic Theories	Purpose
Analyzing the construction of mathematical literacy questions	It consists of 4 contexts (personal, occupational, societal, and scientific), 4 Content domains (quantity, space and shape, change and relationships, and data and uncertainty). It includes 2 main components: the stimulus and the questions.	Golla & Reyes (2022); Stacey (2022); OECD (2021)	Ensures each item comprehensively measures mathematical literacy.
Identification of Real-World Contexts	Selected contexts from daily life (science, society, environment, and technology).	Blum & Leiss (2020); OECD (2021)	Ensures items are meaningful and connected to real situations.
Determination of Competencies	Defined the main competencies: *formulate*, *employ*, and *interpret/evaluate*.	Golla & Reyes, (2022); Tout et al., (2023)	Aligns the instrument with key literacy processes.
Blueprint Design	Create a matrix containing context, content, and competencies, accompanied by Multiple Choice (MC), Complex MC, Short Answer, and Essay question formats.	Mislevy et al.(2021); Wijaya et al.(n.d.)	Become the basis for writing questions consistently
Draft Item Writing	Developed item drafts based on the blueprint, complete with stimulus (text/image), question, and key answer.	Nitko & Brookhart, (2022); Treagust & Chandrasegaran, (2020)	Ensure questions are clear.

Self-Evaluation

The instrument construction process continued by developing each step in Table 2, taking into account the question context, content, cognitive processes, and question form. These results are shown in Table 3.

Table 3. *Distribution of content, context, cognitive processes, and question form*

Content	Cognitive Processes	Contexts	Question Formats
Quantity (5 items)	Formulate, Employ, Interpret-Evaluate	COVID-19 Vaccine Distribution (Societal), Drinking Water Needs (Scientific), Flower Bouquets (Personal)	Short Answer, Multiple Choice (MC), Complex MC, Essay
Change and Relationship (7 items)	Formulate, Employ, Interpret-Evaluate	Sales of Jogja Batik (Occupational), Flower Bouquets (Personal), Drinking Water Needs (Scientific)	Complex MC, Multiple Choice, Matching, Short Answer, Essay
Space and Shape (7 items)	Formulate, Employ, Interpret-Evaluate	Shop Floor Design (Occupational), Gift Boxes (Personal), Crystal Shapes (Scientific)	Short Answer, Complex MC, Matching, Multiple Choice, Essay
Uncertainty and Data (6 items)	Formulate, Employ, Interpret-Evaluate	Olympics 2024 (Societal), Indonesian Temperature Data Analysis (Scientific)	Multiple Choice, Complex MC, Short Answer, Essay

Referring to Table 3. The distribution of content, context, cognitive processes, and question form in the developed literacy instrument, as a self-evaluation phase, comprises 10 stimuli across four contexts, with each stimulus containing 2-3 literacy questions, for a total of 25 math literacy items. The 25 questions used as a draft of the mathematical literacy assessment instrument were rechecked for conformity to the grid before proceeding to the next stage: expert reviews and one-to-one assessments. Figure 3 is an example of a personal context, showing content space and shape, with short-answer questions.

Expert Review and One-to-One Results

The expert review stage assesses the validity of the questions through a logical validation process, while the one-to-one assessment assesses their readability. The results of the expert reviews, which measured the content, construction, language, and norms aspects of the literacy instrument, are shown in Figure 2. Mean Aiken's V ($\bar{V}$) Scores

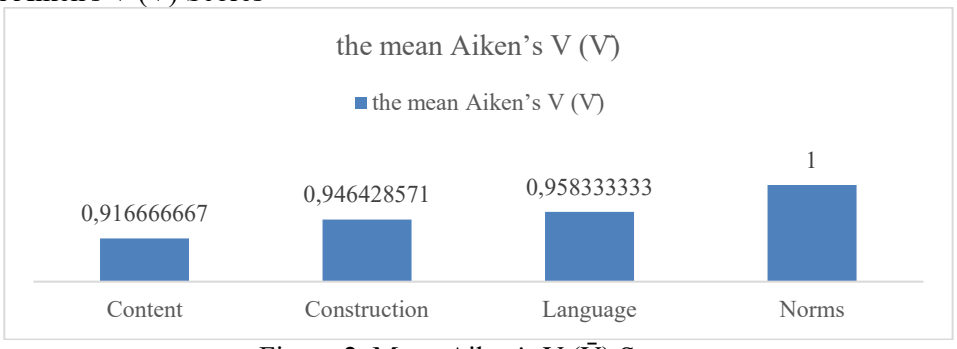


Figure 2. Mean Aiken's V ($\bar{V}$) Scores

The quality of the mathematical literacy instrument had the highest average scores in the construction, language, and norms aspects. Comprehensive content validation across all aspects was calculated using equation 1 and an average validation index $V^{-} = 0.95$, indicating high validity, with only minor revisions needed for wording and clarity. Furthermore, all these values exceeded the rejection threshold ($V^{-} > 0.88$), indicating that the agreement between the two experts was statistically significant, strong, and free from random judgment.

One validator suggested reviewing the answer key for number 11. The initial answer resembles the one used by many teachers in Figure 4, while the revised answer key is shown in Figure 3b. Let us look at the context and problem in Figure 3. Rechecking the answer key is necessary because, mathematically, the area required to make 10 gift boxes is 4 times the area of F4 paper. However, in the context of the

problem, 1 F4 paper can only make 2 boxes, requiring a minimum of 5 F4 papers is required. In this process, a visualization process is required, not just mathematical calculations.

(a) (b)

Figure 3. Context personal (a) and problem number 11 (b)

The one-to-one phase was conducted concurrently with the expert reviews phase. Teachers assessed readability based on the clarity of the test instructions, each stimulus, and the purpose of each question; the communicative wording of the stimulus and questions; the absence of ambiguity; understandable illustrations; and the workability of the questions. The average score obtained from the one-to-one assessment was 4.14, placing the mathematical literacy instrument in the good category for readability. This category requires minor revisions. Revisions were made by clarifying the test instructions for each stimulus. The revised instrument was then subjected to small-group and focus group discussions.

Quality of the Small Group and Focus Group Discussion Instrument Results

All respondents, 14 junior high school mathematics teachers from all regencies/cities in Yogyakarta, completed 25 mathematical literacy questions within 2 hours (80 minutes), meaning the time to complete the questions fit within the available time. After completing the questions, teachers completed a questionnaire via GForm to confirm the questions' difficulty, practicality, and the obstacles or challenges they faced in completing them. Analysis of teacher responses revealed that five questions required more time than others. In addition to these questions, the remaining questions were easy or moderate. After confirming the difficulty index analysis, these five questions fell into the moderate and difficult categories. A more detailed analysis of the difficulty indices for the five questions is presented in Table 4.

Table 4. *Questions requiring more time to complete and their difficulty indices.*

No	Content	Context	Mathematical reasoning	Questions format	Difficulty index	Difficulty Level
10	space and shape	Personal	Employ	complex multiple choice	0,64	Medium
11	space and shape	Personal	Interpret and evaluate	short answer	0,14	Hard
15	change and relationship	Scientific	Employ	complex multiple choice	0,43	Medium
19	space and shape	Occupational	Employ	essay	0,36	Medium
24	space and shape	Scientific	Interpret and evaluate	short answer	0,57	Medium

In Table 4 above, the greatest difficulty was observed for item 11, with a difficulty index of 0.14. The error occurred because the teacher did not consider the question's context and the overall visualization process, as shown in Figure 4. Example of Teacher's Answer. This type of question focuses on short answer questions, focusing on the content of space and shape, and the reasoning level of interpretation and evaluation. Furthermore, teachers found the content difficult, particularly space and shape, which require mathematical reasoning, employing, interpreting, and evaluating. Therefore, teachers should focus

more on developing higher-level reasoning skills. Based on the question format, complex multiple-choice and short-answer questions dominated the question types teachers considered difficult.

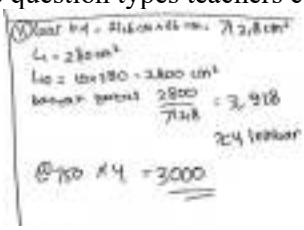


Figure 4. Example of Teacher's Answer

The results of the qualitative analysis of teachers' responses to the questions are shown in Figure 5 below.

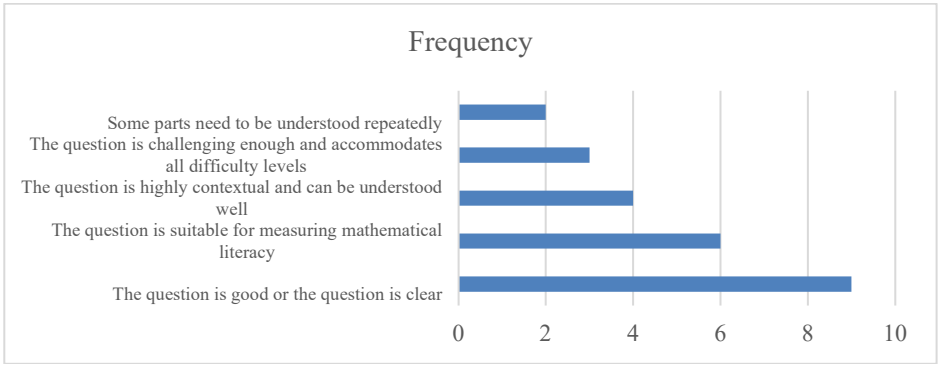


Figure 5. Teacher Response Analysis

The majority of teachers (64.3%) stated that the questions were good and clear. Six teachers stated that the questions were appropriate for measuring literacy, and only two felt they needed to be read repeatedly. The results indicate that the literacy instrument can be used to measure teachers' literacy skills and accommodates all levels of difficulty.

Field Test

The results of the field test on 45 teachers showed that each of the 25 developed questions had a sig value ≤ 0.05, or, when compared to the r table with df = 43 (45-2), the Pearson correlation for each question item was> 0.2940. Therefore, the analysis of the questions shows that each question developed is valid. In addition, the Cronbach's alpha was 0.860, which is> 0.7 (Figure 6). This shows that, overall, the developed question instrument is reliable and consistent. The difficulty index at the field test stage has the same results as the small group stage. From the discriminant power analysis, question number 11 is difficult, as are the easy and medium-level questions. This means there is consistency between the small group and field test results. Therefore, the question instrument is considered reliable. Thus, in addition to meeting the validity criteria for constructs, the developed questionnaire also demonstrates empirical validity.

Figure 6. Reliability Results from *SPSS Output*

DISCUSSION

The process of constructing mathematical literacy instruments for teachers aligns with the principles of educational assessment development (Stacey, Kaye & Turner, 2015). The literacy instrument integrates real-world contexts, namely Flower Bouquets and Gift Boxes (personal context),

the 2024 Olympics and Covid Vaccine Distribution (Societal context), Sales of Jogja Batik and Shop Floor Design (Occupational context), and Drinking Water Needs, Crystal Shapes, and Indonesian Temperature Data Analysis (scientific context). The use of these contexts avoids abstract mathematical structures, transforms them into easily understood contexts, and makes the test items more meaningful (Blum & Leiss, 2020).

The numeracy literacy questions developed not only test procedural skills but also engage a more comprehensive cognitive process, in both content and reasoning. Based on the answer to question 11, teachers need a thinking process that includes formulating, applying, and interpreting/evaluating, as outlined in the PISA framework (Golla & Reyes, 2022b; OECD, 2023; Tout et al., 2023). Although the questions are based on PISA, their contexts are adapted to local settings (Citrawan et al, 2024), such as batik, weather patterns in Indonesia, and other locally relevant data. In contrast, PISA employs broader problem contexts. This constitutes the distinctive feature that sets the questions developed in this study apart from those used in PISA.

The suitability of the construction process is further reflected in the results of expert validation. The average Aiken index ($V^{-} = 0.95$) exceeds the minimum statistical significance threshold of the Aiken Table (0.88). The instrument achieved a perfect score (5) in the content aspect because the stimulus is up-to-date and not multi-interpretable, and the questions can measure literacy skills and have answer options that can be solved. It achieved a maximum score in the construction aspect because the questions align with the content, context, and level of reasoning being measured. This is because the questions were well-structured following the PISA framework. The instrument achieved a maximum score in the norms aspect because the stimuli and questions do not touch on ethnicity, religion, race, or intergroup relations (SARA), politics, or pornography. The language aspect scored lower than the other aspects because several sentences in the stimuli and questions needed to be reviewed to comply with proper and correct EYD. The follow-up process involved revising the questions and stimuli to conform to proper and correct language.

In addition to meeting content validity, the small-group results strengthened the validity findings because teachers stated that the instrument was contextual, addressed current issues, and aligned with the mathematical literacy framework of PISA (OECD, 2021). The question format conforms to the AKM (Academic Skills Standards)(Wijaya et al., n.d.). The questions can vary in difficulty, so the instrument also includes items at different difficulty levels. Many teachers found it challenging to solve contextual problems that demanded reasoning beyond procedural steps. This aligns with previous research showing that teachers need stronger skills in higher-order thinking and contextual problem-solving (OECD, 2021a; Tout et al., 2017; Wijaya et al., 2024).

The process of developing the question instrument from the preliminary stage to the field test stage indicates that the instrument has undergone a complete development cycle. The outcome of this phase was the testing of the instrument on 45 junior high school mathematics teachers. These respondents exceeded the minimum pilot reliability requirement of 30 respondents in the initial Cronbach's alpha/KR-20 test (Suciati et al., 2020; Wong et al., 2023). This research process strengthened the instrument development phase of Benson and Florence Clark (Benson & Clark, 1982), which includes planning, construction, quantitative evaluation, and validation.

Empirical data from both small-group and field tests found that the Space and Shape content domain presents a unique challenge for teachers because it requires challenging spatial reasoning. This is consistent with research by Sinclair et al. (2016), which found that the development of spatial reasoning and geometry is often a major obstacle due to a lack of visualization. This condition occurs not only among newly graduated teachers but also in teacher-training contexts. Educators tend to be accustomed to procedural algebraic geometric proofs and experience difficulties when faced with non-routine conceptual visualization and spatial interpretation tasks.

CONCLUSION

This study produced 25 test items as a valid and reliable quality teacher mathematical literacy assessment instrument. The instrument construction process is in accordance with the principles of developing educational assessment and mathematical literacy in the PISA framework. Each test item is

aligned with a real-world context that facilitates a different reasoning process than routine questions that prioritize procedural knowledge. The test instrument has a high content-validation category, with an average Aiken index of 0.95 (Aiken's $\bar{V} = 0.95$). The instrument has a good readability aspect (mean = 4.14). The field test results show that each test item has a Sig. ≤ 0.05 value in the correlation test, or a correlation coefficient > 0.2940. The Cronbach's Alpha value is 0.860, which is> 0.7, so the instrument is also reliable. These results indicate that the literacy instrument can measure mathematical literacy both theoretically and practically, as it also has a diverse difficulty index. The literacy instrument has great potential to enrich teachers' ability to create literacy questions. This instrument can theoretically serve as an alternative PISA-based mathematical literacy measurement tool that is teacher-oriented, as well as a practical guide for educators in designing innovative, non-procedural classroom assessments.

ACKNOWLEDGMENTS

The researchers would like to express their deepest gratitude to the validators and junior high school mathematics teachers who are members of the MGMP Mathematics forum in 5 districts/cities in Yogyakarta Province and who participated in the face-to-face sessions, small-group tests, and focus group discussions (FGDs). Their support in this research was very constructive.

DECLARATIONS

Author Contribution	:	Niluh Sulistyani: Conceptualization; methodology; formal analysis; writing the original draft. Albertus Hariwangsa Panuluh: Conceptualization; methodology; investigation; review and editing; translation. Kintan Limiansih: validation, writing, resources; data curation; visualization; supervision; project administration. Margaretha Madha Melissa: review and editing; instrument development. All authors have read and approved the published version of the manuscript.
Funding Statement	:	This research was funded by the Institute for Research and Community Service (LPPM), Universitas Sanata Dharma, under the internal research grant scheme, contract No. 016./Penel./LPPM-USD/Iy2025.
Conflict of Interest	:	The authors declare no conflict of interest. All procedures in this study were conducted in accordance with ethical standards for educational research. The research team obtained permission from the Yogyakarta Provincial Education Office and from the Regency/City Education Offices throughout the province. Respondent participation and consent were voluntary. The confidentiality of all respondent data was strictly maintained.
Additional Information	:	Additional information is available for this paper.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*(1), 131–142. https://doi.org/https://doi.org/10.1177/0013164485451012

Aisyah Aisyah, D. J. (2022). The description of Indonesian students' mathematics literacy in the last decade. *International Journal of Trends in Mathematics Education Research*. https://doi.org/https://doi.org/10.33122/ijtmer.v5i1.114.

Alfaruqi, A. Z., & Nurwahidayah. (2025). *Jurnal Pendidikan IPS*. *15*(1), 11–19. https://doi.org/https://doi.org/10.37630/jpi.v15i1.2559

Amin, B., Baharuddin, B., Nur, F., Nursalam, N., & Angriani, A. D. (2025). Pengembangan Tes Asesmen Kompetensi Mahasiswa Untuk Mengukur Kemampuan Literasi Numerasi. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *14*(1), 54. https://doi.org/10.24127/ajpm.v14i1.9494

Apriandi;, D., Murtafiah;, W., Suprapto;, E., Masfingatin;, T., & Lusiana, R. (2024). The development of problem-based calculus E-modules to improve students' numeracy literacy skills. *AIP Conference Proceedings*, *3148*(1). https://doi.org/https://doi.org/10.1063/5.0243257

Benson, J., & Clark, F. (1982). A Guide for Instrument Development and Validation. *The American Journal of Occupational Therapy*, *36*(12), 789–800. https://doi.org/https://doi.org/10.5014/ajot.36.12.789

Blum, W., & Leiss, D. (2020a). The Role of Contexts in Mathematical Modeling – Towards a Conceptual Framework. *ZDM Mathematics Education*, *52*, 149–163. https://doi.org/10.1007/s11858-019-01100-1

Chen, X., Zhou, J., Wang, J., Wang, D., Liu, J., Shi, D., Yang, D., & Pan, Q. (2022). Visualizing Status, Hotspots, and Future Trends in Mathematical Literacy Research via Knowledge Graph. *Sustainability (Switzerland)*, *14*(21). https://doi.org/10.3390/su142113842

Chiu, T. K. F., Chen, Y., Yau, K. W., Chai, C. Sing, Meng, H., King, I., Wong, S., & Yam, Y. (2024). Developing and validating measures for AI literacy tests: From self-reported to objective measures. *Computers and Education: Artificial Intelligence*, *7*(August), 100282. https://doi.org/10.1016/j.caeai.2024.100282

Citrawan, I. W., et al. (2024). *Special Education Teachers ' Ability in Literacy and Numeracy Assessments Based on Local Wisdom*. *8*(1), 145–157. https://doi.org/https://doi.org/10.22437/jiituj.v8i1.32608

DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (Second Edition). SAGE Publications.

Golla, E., & Reyes, V. (2022b). Developing Mathematical Literacy Test Items for Junior High School Teachers Based on the PISA Framework. *International Journal of Evaluation and Research in Education*, *11*(4), 2045–2058. https://doi.org/10.11591/ijere.v11i4.21983

Helstad, K., Solbrekke, T. D., & Wittek, A. L. (2017). Exploring teaching academic literacy in mathematics in teacher education. *Education Inquiry*, *8*(4), 318–336. https://doi.org/10.1080/20004508.2017.1389225

Hilmi, I. &
Kismiantini. (2024). Teacher Competencies, The Shortage of School Resources, and Mathematics Achievement Based on PISA 2018 Indonesia. *Jurnal Pendidikan Matematika*, *18*(2). https://doi.org/https://doi.org/10.22342/jpm.v18i2.pp245-258

Holenstein, M., Bruckmaier, G., & Grob, A. (2021). Transfer effects of mathematical literacy: an integrative longitudinal study. *European Journal of Psychology of Education*, *36*(3), 799–825. https://doi.org/10.1007/s10212-020-00491-4

Hwang, J., & Ham, Y. (2021). Relationship between mathematical literacy and opportunity to learn with different types of mathematical tasks. *Journal on Mathematics Education*, *12*(2), 199–222. https://doi.org/10.22342/JME.12.2.13625.199-222

Kandil, S., & Işıksal-Bostan, M. (2019). Effect of inquiry-based instruction enriched with origami activities on achievement and self-efficacy in geometry. *International Journal of Mathematical Education in Science and Technology*, *50*(4), 557–576. https://doi.org/10.1080/0020739X.2018.1527407

Kappassova, S., Abylkassymova, A., Bulut, U., Zykrina, S., Zhumagulova, Z., & Balta, N. (2025). Mathematical literacy and its influencing factors: A decade of research findings (2015-2024). *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(7). https://doi.org/10.29333/ejmste/16615

Kioupi, V., & Voulvoulis, N. (2019). Education for sustainable development: A systemic framework for connecting the SDGs to educational outcomes. *Sustainability (Switzerland)*, *11*(21). https://doi.org/10.3390/su11216104

Manfreda Kolar, V.M, Hodnik, T. (2021). Mathematical Literacy from the Perspective of Solving Contextual Problems. *European Journal of Educational Research*, *10*(1), 467–483. https://doi.org/10.12973/eu-jer.10.1.467

McKay, V. (2018). Literacy, lifelong learning, and sustainable development. *Australian Journal of Adult Learning*, *58*(3), 390–425. https://ala.asn.au/wp-content/uploads/2021/01/AJAL-58-3-Full.pdf

McKenney, S., & Reeves, T. C. (2018). About educational design research. *Conducting Educational Design Research*, 5–33. https://doi.org/10.4324/9781315105642-3

Merino-Soto, C. (2023). Coeficientes V de Aiken: diferencias en los juicios de validez de contenido. *MHSalud: Revista en Ciencias del Movimiento Humano y Salud*, *20*(1), 1–10. https://doi.org/10.15359/mhs.20-1.3

Miles, M. B.; Huberman, M. S. J. (2014). *Qualitative Data Analysis* (3rd ed.). SAGE Publication.

Mislevy, R. J., Haertel, G., & Almond, R. G. (2021). Evidence-Centered Design for Educational Assessment: A Research Update. *Educational Measurement: Issues and Practice*, *40*(1), 16–31. https://doi.org/10.1111/emip.12424

Muhaimin, L. H., Sholikhakh, R. A., Yulianti, S., Ardani, H., A., & Sahara, S. (2024). Unlocking the secrets of students' mathematical literacy to solve mathematical problems: A systematic literature review. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(4). https://doi.org/10.29333/ejmste/14404

Nadila, S. A., & Lestiana, H. T. (2024). Kemampuan Literasi Matematika Siswa Berdasarkan Gaya Belajar. *Edumatica: Jurnal Pendidikan Matematika*, *14*(03). https://doi.org/https://doi.org/10.22437/edumatica.v14i3.38303

Nieveen, N., & Folmer, E. (2013). Formative Evaluation in Educational Design Research. In *Educational Design Research – Part A: An Introduction* (pp. 152–169). Netherlands Institute for Curriculum Development (SLO).

Nitko, A. J., & Brookhart, S. M. (2022). *Educational Assessment of Students* (8th ed.). Pearson Education.

OECD. (2016). *Program for International Student Assessment (PISA) Results From PISA 2015*. https://www.oecd.org/pisa/PISA-2015-Indonesia.pdf

OECD. (2019). *PISA 2018 Results*. https://www.oecd.org/pisa/Combined_Executive_Summaries_PISA_2018.pdf

OECD. (2021). *PISA 2021 Mathematics Framework*. https://www.oecd.org/pisa/pisaproducts/pisa-2021-mathematics-framework.pdf

OECD. (2023). PISA 2022 Assessment and Analytical Framework. In *Paris: Journal of OECH Publishing*.

PISA. (2023). PISA 2022 Results Factsheets Indonesia. *The Language of Science Education*, *1*, 1–9. https://oecdch.art/a40de1dbaf/C108.

Pulsiri, N., Vatananan-Thesenvitz, R., Tantipisitkul, K., Aung, T. H., Schaller, A. A., Schaller, A. M., Methananthakul, K., & Shannon, R. (2019). Achieving sustainable development goals for people with disabilities through digital technologies. *PICMET 2019 - Portland International Conference on Management of Engineering and Technology: Technology Management in the World of Intelligent Systems, Proceedings*, (November 2020). https://doi.org/10.23919/PICMET.2019.8893725

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

Sinclair, N., Bartolini Bussi, M. G., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: an ICME-13 survey team report. *ZDM*, *48*(5), 691–719. https://doi.org/10.1007/s11858-016-0796-6

Stacey, K. (2011). The PISA view of mathematical literacy in Indonesia. *Journal on Mathematics Education*, *2*(2), 95–126. https://doi.org/10.22342/jme.2.2.746.95-126

Stacey, K. & Turner, R. (2015). *Assessing Mathematical Literacy: The PISA Experience* (R. Stacey, Kaye & Turner, Ed.). Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-319-10121-7

Suciati, S. Munadi, Sugiman, & Wiwin Dwi Ratna Febriyanti. (2020). Design and Validation of Mathematical Literacy Instruments for Assessment for Learning in Indonesia. *European Journal of Educational Research*, *9*(2), 865–875. https://doi.org/https://doi.org/10.12973/eu-jer.9.2.865

Sugiyono. (2019). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D*. Alfabeta.

Tessmer, M. (2005). *Planning and Conducting Formative Evaluations: Improving the Quality of Education and Training*. Routledge.

Tout, D., Gal, I., & Schlöglmann, W. (2023). Mathematical Reasoning and Literacy: Revisiting PISA's Conceptualization for the Next Decade. *Studies in Educational Evaluation*, *78*, 102210. https://doi.org/10.1016/j.stueduc.2023.102210

Tout, D., Coben, D., Geiger, V., Ginsburg, L., Hoogland, K., Maguire, T., Thomson, S., & Turner, Ross. (2017). *Review of the PIAAC Numeracy Assessment Framework: Final Report.*

Treagust, D. F., & Chandrasegaran, A. L. (2020). Developing Diagnostic Assessment Items to Evaluate Students' Understanding of Science and Mathematics Concepts. *International Journal of Science and Mathematics Education*, *18*, 689–708. https://doi.org/10.1007/s10763-019-09971-0

Umbara, U., & Suryadi, D. (2019). Re-interpretation of mathematical literacy based on the teacher's perspective. *International Journal of Instruction*, *12*(4), 789–806. https://doi.org/10.29333/iji.2019.12450a

UNESCO. (2024). Global report on teachers: addressing teacher shortages and transforming the profession. In the *Global report on teachers: addressing teacher shortages and transforming the profession*. https://doi.org/10.54675/figu8035

Wijaya, A., Retnawati, H., & Nusantara, T. (2024). Enhancing Teachers' Competence in Developing Mathematical Literacy Assessment Aligned with the PISA Framework. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(2), em2295. https://doi.org/10.29333/ejmste/13670

Wijaya, A., Yogyakarta, U. N., Dewayani, S., Litara, Y., Effendi, A., Yogyakarta, U. N., & Gunawan, H. (n.d.). *Framework Asesmen Kompetensi Minimum (AKM)*.

Wong, V., Yun, S., Ulang, N., & Husain, S. H. (2023). *Measuring the Internal Consistency and Reliability of the Hierarchy of Controls in Preventing Infectious Diseases on Construction Sites : The Kuder-Richardson (KR-20) and Cronbach's Alpha*. *1*(1), 392–405. https://doi.org/https://doi.org/10.37934/araset.33.1.392405

Yusmar, F., & Fadilah, R. E. (2023). Analisis Rendahnya Literasi Sains Peserta Didik Indonesia: Hasil Pisa Dan Faktor Penyebab. *LENSA (Lentera Sains): Jurnal Pendidikan IPA*, *13*(1), 11–19. https://doi.org/10.24929/lensa.v13i1.283

