NUMERACY LITERACY OF GRADE VII STUDENTS IN SOLVING CONTEXTUAL PROBLEMS ACROSS MATHEMATICAL ABILITY LEVELS

Hikmatul Fadhilah Sianipar[1], Izwita Dewi[2], Kms. Muhammad Amin Fauzi[3]
[1]Departement of Electrical Engineering, Universitas Pembangunan Panca Budi,
[2,3]Departement of Mathematics Education, Universitas Negeri Medan
e-mail: fadhilah525@gmail.com

Abstract
Numeracy literacy is a fundamental competency assessed through Indonesia's Minimum Competency Assessment (AKM); however, Grade VII junior high school students' achievement in this domain continues to exhibit significant gaps, particularly in solving contextual problems with extended narrative structures. Although numeracy literacy research in Indonesia has grown considerably, studies that simultaneously examine students' profiles across differentiated ability levels while foregrounding the linguistic dimension as the primary explanatory variable remain scarce—especially in the local context of Stabat. To address this gap, this study investigated the numeracy literacy profile of Grade VII students at a junior high school in Stabat across three levels of mathematical ability: high, medium, and low. A descriptive qualitative approach was employed involving 20 students selected through purposive sampling. Instruments consisted of a contextual numeracy literacy test and a semi-structured interview guide. Data analysis followed the Miles and Huberman interactive model. Findings indicate that 60% of students fall within the medium ability level. More critically, the primary barrier originates from linguistic demands rather than procedural weaknesses: students demonstrate decreased accuracy when confronted with word problems exceeding three sentences, and the capacity for decoding—translating verbal context into a mathematical model—remains underdeveloped across all ability levels. The principal contribution of this study is establishing that linguistic load, not computational deficiency, is the decisive differentiating factor in numeracy literacy performance, with direct pedagogical implications: strengthening numeracy literacy cannot be decoupled from simultaneously developing students' reading stamina and mathematical reading strategies.
Keyword: *Numeracy Literacy, Contextual Problems, Mathematical Ability Levels, Linguistic Aspects, AKM*

INTRODUCTION

Since the introduction of Indonesia's Minimum Competency Assessment (AKM) in 2021, numeracy literacy has emerged as a central competency in national education — defined not merely as arithmetic proficiency, but as the capacity to use numbers and mathematical reasoning to solve practical, context-rich problems across diverse life situations (Ministry of Education and Culture, 2020, 2021). This reconceptualization is consequential: it positions authentic mathematical activity at the intersection of quantitative reasoning and language comprehension, making linguistic ability an inherent component of numeracy performance rather than a peripheral support skill.

PISA's three-process framework — formulating, employing, and interpreting mathematical situations — confirms that numeracy literacy is inherently language-dependent: students must decode verbal context before any mathematical procedure can begin (OECD, 2022). Indonesia's standing in PISA 2022 (70th out of 81 countries; average score 366 versus the OECD mean of 472) makes this linguistic-mathematical gap empirically urgent, pointing to a systemic overemphasis on procedural computation at the expense of contextual reasoning (OECD, 2023; Sholihah & Susanti, 2023).

Research consistently identifies linguistic load — the cognitive burden imposed by the textual complexity of word problems — as a primary barrier to numeracy performance, distinct from and often independent of procedural weakness (Abedi & Lord, 2001; Kolar & Hodnik, 2021; Verschaffel et al., 2020). Narrative length is particularly consequential: as problem text grows longer, students increasingly adopt a counterproductive number-hunting strategy — bypassing contextual reading to extract numerical values directly — which systematically undermines the reasoning that numeracy tasks require (Daroczy et al., 2015; Kirkland & McNeil, 2021; Wijaya et al., 2021). Preliminary observations at a junior high school in Stabat revealed a consistent pattern: Grade VII students showed marked confusion when word problems exceeded three sentences, independent of their general computational ability. Despite a growing body of Indonesian numeracy literacy research (Aisyah & Juandi, 2022; Hapsari et al., 2022), no study has simultaneously profiled students across differentiated ability levels while positioning linguistic load as the primary explanatory variable — particularly within the Stabat context. This gap motivates the present study, which aims to: (1) map the numeracy literacy profile of Grade VII students in Stabat across three mathematical ability levels, and (2) identify the specific patterns through which linguistic barriers differentiate performance across those levels — thereby generating actionable evidence for instructional design that integrates linguistic and mathematical competency development.

METHOD

This study employs a qualitative descriptive approach to portray phenomena in depth within their natural context (Cohen et al., 2018; Creswell & Creswell, 2018). This design was selected because the research questions centre on understanding students' thinking patterns and cognitive barriers — phenomena that require triangulation across test data, classroom observation, and semi-structured interviews to be adequately captured (Miles et al., 2014). Trustworthiness was ensured through four Lincoln and Guba (1985) criteria: credibility (via source triangulation and member checking with focal subjects), transferability (via thick contextual description of the research site and participants), dependability (via an audit trail documenting data reduction decisions), and confirmability (via explicit documentation of the inferential chain from raw data to conclusions).

The research was conducted at a State Junior High School in Stabat, Langkat Regency, North Sumatra. A total of 20 Grade VII students were selected through purposive sampling — a strategy appropriate for qualitative inquiry where the goal is informational richness rather than statistical representativeness (Creswell & Creswell, 2018); this sample size is consistent with the range of 15–25 participants recommended for descriptive qualitative studies of this scope (Cohen et al., 2018). Students were grouped by initial diagnostic test scores: High (≥75), Medium (50–74), and Low (<50). Three focal subjects — one per ability level — were then selected for in-depth analysis on the basis of the analytical clarity of their written responses and availability for interview, ensuring rich representation of each ability profile.

The primary instrument was a three-item numeracy literacy test consisting of contextual word problems aligned to the AKM and PISA process dimensions, with deliberate variation in narrative length to expose differential responses to linguistic load. Content validity was established through expert review by two mathematics education specialists and one experienced classroom teacher, who independently assessed item alignment, linguistic appropriateness, and cognitive level; all items were revised until inter-rater consensus was achieved. The secondary instrument was a semi-structured interview guide designed to surface reasoning processes not visible in written responses. Test scores were computed as: Value = (Total Score / 36) × 100, applying the same ability-level cut-offs used in initial diagnostic scoring. Data analysis followed the interactive model of Miles et al. (2014) in three iterative stages: (1) data condensation — open coding of written responses and interview transcripts to identify recurring error types and linguistic barrier patterns; (2) data display — organising coded themes into cross-ability comparison matrices to enable systematic comparison across the three focal subjects; and (3) conclusion drawing — developing interpretive claims verified through source triangulation (test, observation, interview) and member checking with focal subjects to confirm accuracy of interpretations.

Tabel 1 Indicators and Instruments for the Numeracy Literacy Test

No.	Numeracy Literacy Indicators	Question Items / Test Instruments
1	Ability to Formulate Situations Mathematically (Formulating)	Question 1: Mr. Harun sells rubber at Stabat Market at a price of IDR 15,000 per kg. Today he received a total of Rp300,000, including a fixed bonus of Rp45,000 from collectors. Model a one-variable linear equation for this problem (suppose x = many kg of ru
2	Ability to Apply Mathematical Concepts, Procedures, Facts, and Tools (Employing)	Question 2: Mrs. Fatimah has a banana chips business on Jalan Sudirman, Stabat. The production cost is fixed per day of IDR 150,000 and the cost of raw materials is IDR 10,000 per pack. Chips are sold for IDR 35,000 per pack. On Saturday, Mrs. Fatimah sold x packs and made a net profit of IDR 400,000. Determine the value of x!
3	Ability to Interpret and Evaluate Mathematical Results (Interpreting)	Question 3: In Banyumas Village, residents work together to clean irrigation canals. Pak Ali cleaned (2n+5) meters, Pak Bahar (4n−3) meters, and Pak Candra 18 meters. The total channel is 62 meters. (a) Make a one-variable linear equation! (b) If 6 meters of the channel is missed, how does your equation change?

Tabel 2 Numeracy Literacy Assessment Rubric

Indicator	Criteria	Scor
Ability to Formulate Situations Mathematically	Able to identify all relevant information and formulate the situation into a mathematical model accurately and completely without errors.	4
	Able to identify most of the information; The formulation of a good model even though there is one small mistake that is not principled.	3
	Able to identify partial but incomplete information; The formulation of the model is quite adequate but there are some errors.	2
	Lack of ability to identify relevant information; The formulation of the model is inaccurate or not carried out.	1
Ability to Apply Mathematical Concepts, Procedures, Facts, and Tools	Able to apply precise mathematical concepts, procedures, and facts; The steps of solving the process are sequential, systematic, resulting in the correct answer.	4
	Able to implement concepts and procedures well; The steps are mostly correct despite one minor computational error.	3
	Able to apply some concepts and procedures, but there are some errors in the process.	2
	Lack of ability to implement concepts and procedures; Steps are not systematic, many mistakes, or	1

Indicator	Criteria	Scor
	solutions are not made.	
Ability to Interpret and Evaluate Mathematical Results	Able to interpret results into the context of the question appropriately; able to evaluate reasonableness and communicate conclusions completely.	4
	Able to interpret results well; The conclusion is adequate even if there are minor inaccuracies in the unit or completeness of the sentence.	3
	Able to interpret results but there are some errors in interpretation; The conclusion is incomplete or not entirely relevant.	2
	Lack of ability to interpret results; not demonstrating the ability to evaluate fairness; The conclusion is unwritten or irrelevant.	1

RESULT AND DISCUSSION

Distribution of Students' Numeracy Literacy Ability Levels

Tabel 3: Recapitulation of the Distribution of Numeracy Literacy Ability Level for 20

Ability Level	Value Range	Number of Students	Percentage
Height	75 – 100	4 students	20%
Medium	50 – 74	12 students	60%
Low	< 50	4 students	20%
Total		20 students	100%

Table 3 reveals an ability distribution concentrated in the medium category: 12 students (60%) at medium, with symmetrical tails of four students (20%) at each extreme. This pattern replicates findings reported for Indonesian junior high school students nationally — procedural competence is relatively widespread, but contextual reasoning remains a persistent ceiling (Aisyah & Juandi, 2022; Hapsari et al., 2022). Conceptually, this distribution suggests that the educational emphasis on algorithmic mastery has been partially successful, but has not extended to the higher-order demand of translating real-world context into mathematical structure — precisely the transition point where linguistic load becomes decisive. The analytical focus therefore shifts from the distribution itself to the qualitative patterns of error that that explain why the majority cluster at the medium band rather than progressing to high-level performance.

In-depth analysis of three representative subjects: High-Level Subject (S-T), S-T is a student with a final score of 83.3 who completes all three questions sequentially and systematically. In Question 1 (the context of Pak Harun's rubber), S-T wrote down the components known and asked before building the model 15,000x + 45,000 = 300,000 and completed exactly to x = 17 kg. In Question 2, S-T builds the concept of net profit as the difference between revenue and total cost: 35,000x − (150,000 + 10,000x) = 400,000, obtaining x = 22 packs. In Question 3 which has the longest narrative, S-T is still able to identify all components accurately and logically modify the equation in sub-question (b).

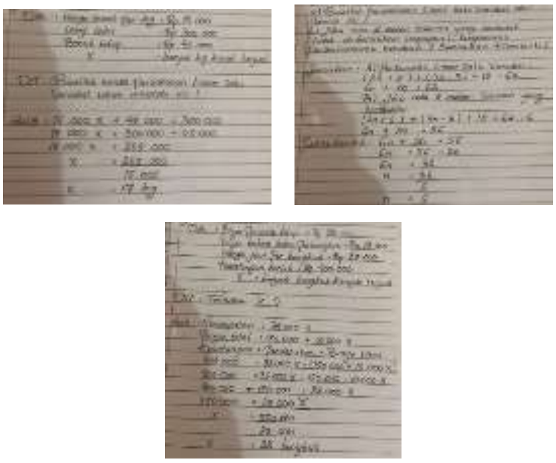

Figure 1. Example of a high-level subject student response

The advantages of S-T are most prominent in its formulation capabilities. When confronted with the long narrative of Question 3, S-T immediately extracts the relational structure from the narrative — reflecting what Kolar and Hodnik (2021) call mathematical reading: reading texts is not just to understand literal meaning, but to identify hidden mathematical structures. In the interpreting aspect, S-T consistently ends the answer with a meaningful contextual statement, showing an understanding that the resulting number is the answer to a contextual question that must be communicated again. *Researcher: "When you read Question 3 which is quite long, what is the first thing you do?" S-T: "I'll read it all to the end, ma'am. As I read, I underlined the important numbers and who worked on how much. Only then do I write what is known and questioned."*

The S-T response confirmed that structured reading strategies are key to adaptation to a high linguistic load. Active reading accompanied by the selection and marking of relevant information (Wijaya et al., 2021) directly supports strong formulating skills and minimizes the risk of misinterpretation.

Medium Level Subject (S-S), S-S is a student with a final score of 63.9 which shows a typical profile of the medium category: able to solve short to medium narrative problems, but experiencing a significant decline in performance in long narrative questions. In Question 1, S-S managed to construct a precise model of the equation and obtained x = 17 kg, albeit without an explicit known-asked stage. In Question 2, S-S ignored the x variable in the income component due to not carefully reading that Mrs. Fatimah was selling "x wraps" — a mistake that stemmed from an incomplete decoding process. In Question 3 sub-question (b), S-S mistakenly interprets "6 meters of missed channels" as an addition (62 + 6 = 68) instead of a subtraction (62 − 6 = 56) — a failure of linguistic inference, not a computational weakness.

Researcher: "Usually if the question is long, do you read everything or just look up the numbers?" S-S: "If it's long, usually look for the number right away, ma'am. If you read everything, you will be confused and forget the first one."

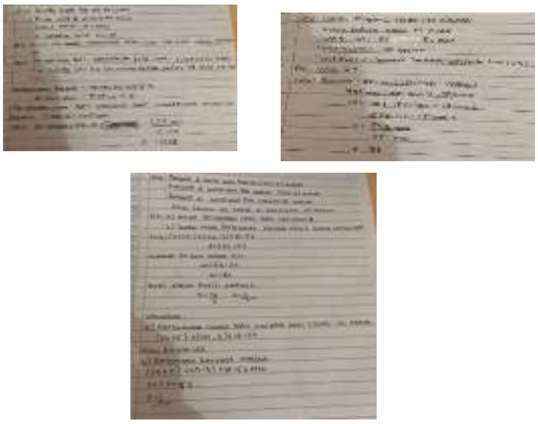

Figure 2. Example of a medium-level subject student response

The number-hunting strategy — directly looking for numerical values without reading the entire narrative — is an adaptive mechanism that is precisely counterproductive because it eliminates the essential context for building proper mathematical models (Hapsari et al., 2022). Medium-level students are susceptible to linguistic ambiguity, particularly in the construction of passive sentences or negative clauses. These findings are in line with Kirkland and McNeil (2021) that the way change conditions are communicated linguistically significantly affects students' sense-making abilities. Low-Level Subject (S-R), S-R is a student with a final score of 38.9 that indicates a double barrier: linguistic and procedural-mathematical. In Question 1, S-R is able to write down the known components completely but only writes 15,000x+45,000 = 300,000 without including the bonus component, indicating a failure to integrate information into a coherent mathematical representation even though the information has been identified. In Question 2, S-R only wrote a few lines of inconsistent equations — an indication of algorithmic imitation without conceptual understanding (Lithner, 2003). In Question 3, S-R only managed to copy the information into a known format without being able to construct any equations; Sub-question (b) is not worked on at all. *Researcher: "For Question 3, why is part (b) not done?" S-R: "I've read it many times, ma'am, but I don't understand what '6 meters missed' means. So I left."*

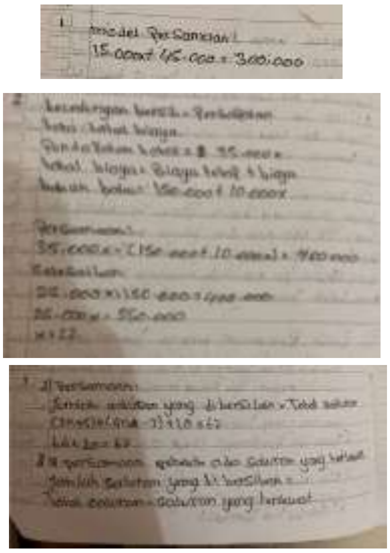

Figure 3. Example of a low-level subject student response

The S-R profile reveals the complex interaction of two barriers that exacerbate each other. Linguistic barriers are not just the inability to read long narratives, but the inability to construct situational representations even from medium narratives. Procedural barriers are related to weak strategy flexibility and low metacognitive awareness: errors at the beginning are not detected and accumulate until the final answer (Schoenfeld, 2016). Main Barrier Patterns: Linguistic Dimensions as Determining Factors**,** Cross-subject synthesis yields a theoretically coherent finding: linguistic processing capacity, not procedural mastery, is the primary differentiator of numeracy literacy performance across ability levels. The contrast between S-T, S-S, and S-R is most pronounced not in computational skill — which all three subjects demonstrate at a basic level — but in the capacity to construct a situational model from a verbal problem, a prerequisite for any subsequent mathematical procedure. This finding is explicable through Sweller’s (2011) cognitive load theory: when the linguistic complexity of a problem exhausts a student's language-processing capacity, working memory resources available for mathematical reasoning are depleted, triggering compensatory strategies such as number-hunting. Critically, this dynamic is moderated by narrative length, sentence complexity, and the degree of implicitness in the problem's conditional structure — each of which increases extraneous cognitive load independently of the underlying mathematical difficulty (Daroczy et al., 2015; Verschaffel et al., 2020).

Table 4. Comparison of Numeracy Literacy Barriers Pattern Based on Ability Level

Dimensions of Analysis	High Level (S-T)	Medium Level (S-S)	Low Level (S-R)
Reading Strategy	Thorough and structured reading; Mark Active Key Information	Partial skimming; Instantly Number - Hunting	Repeat reading without selection strategies; Giving Up on Long Narratives
Decoding Capabilities	Accurately extracting relational structures from long narratives	Capable of short narratives; decoding failure on long/implicit narratives	Decoding failure even on medium narratives; Failing to build a situation model
Formulating Ability	Consistently build accurate and complete mathematical models	Right on a simple question; significant errors in complex problems	Often fail to build a model; Information components are not integrated
Linguistic Load Response	Stable performance; No significant decline	A noticeable decrease in accuracy; Increased misinterpretation	Drastic decline; tend not to do problems
Metacognitive Awareness	Actively monitor; Recognize and correct your own mistakes	Limited monitoring; does not always detect answer inconsistencies	Monitoring is almost non-existent; Initial errors accumulated

Table 4 makes visible what aggregate scores conceal: the three-level comparison exposes a gradient not of mathematical knowledge, but of language-mediated reasoning capacity. Numeracy literacy emerges from this table as a tripartite construct — an interaction of linguistic processing, mathematical procedure, and metacognitive self-regulation — with each component necessary but none individually sufficient for high-level performance (Schoenfeld, 2016; Kolar & Hodnik, 2021). These findings extend Wijaya et al. (2021) by specifying that verbal-to-mathematical transformation failure is not uniform: it is calibrated to narrative length, with performance declining predictably as linguistic demand increases, independent of the mathematical complexity of the underlying task. This distinction carries direct pedagogical significance: remediation strategies targeting only procedural fluency will systematically fail students whose core barrier is decoding rather than computation.

Implications for Math Learning, The findings of this study are broadly consistent with international evidence linking linguistic load to mathematics performance gaps. Studies conducted in diverse contexts — including Turkish secondary school students (Daroczy et al., 2015) and multilingual classrooms across Europe (Verschaffel et al., 2020) — similarly identify the verbal-to-mathematical translation stage as the primary locus of failure, particularly when narrative length and grammatical complexity increase simultaneously. This convergence strengthens the generalisability of the present findings beyond the Stabat context. However, a critical caveat must be acknowledged: linguistic load is not the sole explanatory factor. Socio-economic background, reading habits formed outside school, and prior exposure to narrative-rich mathematical tasks represent plausible confounding variables that this study — with its single-school, cross-sectional design — cannot disentangle. Future research should employ controlled designs to isolate the independent contribution of each factor.

Based on these findings, three pedagogical recommendations can be formulated. First, the integration of mathematical reading strategies into regular learning — teachers need to explicitly teach structured text-decoding strategies (annotating known quantities, paraphrasing conditions, diagramming relational structure) as a pre-modelling routine before initiating mathematical procedures. This strategy gradually builds the mathematical reading stamina needed for AKM and has demonstrated effectiveness in reducing linguistic load in comparable contexts (Verschaffel et al., 2020). Second, the design of practice questions with scaffolded linguistic complexity — progressively increasing narrative length from single-sentence to multi-clause problems — allows students to develop linguistic processing capacity in parallel with procedural fluency (Larmer et al., 2015). Third, strengthening metacognitive awareness through think-aloud practices and problem-solving journals, given that metacognition is the strongest independent predictor of mathematical problem-solving success across ability levels (Schoenfeld, 2016) and can be developed through intentional classroom practices.

CONCLUSION

This study has investigated the numeracy literacy profile of 20 Grade VII junior high school students in Stabat across three differentiated ability levels, revealing linguistic load as the primary determinant of contextual problem-solving performance. Three conclusions merit emphasis. First, the concentration of students (60%) in the medium ability band reflects a structural feature of mathematics education that has successfully cultivated procedural competence but has not extended it to contextual reasoning — the transition point where linguistic demand becomes decisive. Second, linguistic barrier — particularly the cognitive overload triggered by narratives exceeding three sentences — functions as the primary differentiator across ability levels, operating independently of procedural competence and manifesting specifically at the decoding stage of mathematical modelling. Third, the number-hunting strategy observed in medium and low-level students is not a random error but a systematic cognitive adaptation to linguistic overload, which must be addressed at the level of instructional design rather than individual remediation. Theoretically, this study contributes to the reconceptualization of numeracy literacy as a tripartite construct — integrating linguistic processing, mathematical procedure, and metacognitive regulation — thereby extending the largely unidimensional view of numeracy implicit in current AKM assessment frameworks and adding local empirical grounding to international cognitive load research (Sweller, 2011; Verschaffel et al., 2020). Practically, this finding motivates three specific, immediately implementable changes: (1) the integration of structured text-decoding instruction (annotation, paraphrasing, relational diagramming) as a pre-modelling routine in every mathematics lesson involving word problems; (2) the deliberate sequencing of word problems by narrative complexity — not only by mathematical difficulty — in curriculum and assessment design; and (3) the adoption of think-aloud protocols as a routine diagnostic tool to make students' linguistic processing visible to teachers, enabling targeted scaffolding before the mathematical modelling phase begins.

The study had several limitations: the limited number of subjects from one school limited generalization; one-session data collection cannot capture the dynamics of longitudinal development; and contextual factors such as socio-economic conditions and reading habits outside of school are not explicitly controlled. For further research, it is recommended: (1) larger-scale research covering several schools in Langkat Regency to verify the pattern of findings; (2) experimental research testing the effectiveness of scaffolded linguistic complexity interventions; and (3) longitudinal research tracing the development of mathematical reading stamina from grades VII to IX.

REFERENCES

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. Applied Measurement in Education, 14(3), 219–234. https://doi.org/10.1207/S15324818AME1403_2

Aisyah, A., & Juandi, D. (2022). The description of Indonesian student mathematics literacy in the last decade. International Journal of Trends in Mathematics Education Research, 5(1), 105–110. https://doi.org/10.33122/ijtmer.v5i1.114

Cohen, L., Manion, L., & Morrison, K. (2018). Research methods in education (8th ed.). Routledge.

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.

Hapsari, I. P., Saputro, T. V. D., & Sadewo, Y. D. (2022). Mathematical literacy profile of elementary school students in Indonesia: A scoping review. Journal of Educational Learning and Innovation (ELIa), 2(2), 213–225. https://doi.org/10.46229/elia.v2i2.513

Himmi, N., Armanto, D., Dewi, I., Surya, E., & Syahputra, E. (2024). Investigation of students' mathematical mindset in solving mathematical literacy problems. AXIOM: Jurnal Pendidikan dan Matematika, 13(2), 139–152. https://doi.org/10.30821/axiom.v13i2.20440

Kementerian Pendidikan dan Kebudayaan. (2020). Panduan Asesmen Kompetensi Minimum. Pusat Asesmen dan Pembelajaran, Kemdikbud.

Kementerian Pendidikan dan Kebudayaan. (2021). Merdeka Belajar episode 17: Asesmen Nasional. Kemdikbud RI.

Kirkland, P. K., & McNeil, N. M. (2021). Question design affects students' sense-making on mathematics word problems. Cognitive Science, 45(4), e12960. https://doi.org/10.1111/cogs.12960

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

Larmer, J., Mergendoller, J. R., & Boss, S. (2015). Setting the standard for project based learning. ASCD.

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. Educational Studies in Mathematics, 52(1), 29–55. https://doi.org/10.1023/A:1023683716659

Manggala, I. S. A., & Yuniawatika. (2021). Develop students' mathematical literacy in learning. Advances in Social Science, Education and Humanities Research, 565, 315–320. https://doi.org/10.2991/assehr.k.211210.044

Melisa, M., & Utami, N. S. (2023). Students' mathematical literacy in solving problems equation of three variables. Jurnal Pendidikan Matematika (JUPITEK), 6(1), 42–48. https://doi.org/10.30598/jupitekvol6iss1pp42-48

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE Publications.

OECD. (2022). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

OECD. (2023). PISA 2022 results (Volume II): Learning during and from disruption. OECD Publishing. https://doi.org/10.1787/a97db61c-en

Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. Journal of Education, 196(2), 1–38. https://doi.org/10.1177/002205741619600202

Sholihah, U., & Susanti, V. D. (2023). Eliciting activities model on students' mathematical literacy ability. AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 12(1), 134–144. https://doi.org/10.24127/ajpm.v12i1.6608

Sweller, J. (2011). Cognitive load theory. Psychology of Learning and Motivation, 55, 37–76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Wijaya, A., Elmaini, & Doorman, M. (2021). A learning trajectory for probability: A case of game-based learning. Journal on Mathematics Education, 12(1), 1–16. https://doi.org/10.22342/JME.12.1.12836.1-16

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H. C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. Frontiers in Psychology, 6, Article 348. https://doi.org/10.3389/fpsyg.2015.00348

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. ZDM – Mathematics Education, 52(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4