Journal of Research in Instructional
e-ISSN: 2776-222X
Vol. 6(1) 2026, pp. 438 – 453
https://doi.org/10.30862/jri.v6i1.93

Integration of socio-cultural digital teaching materials to enhance problem-solving and self-efficacy in elementary school mathematics

Nurlaela Nurlaela[1,*], Slamte Arifin, Siti Faizah

Universitas Negeri Malang, Indonesia

Submitted: 25-09-2025

Accepted: 20-02-2026

Published: 30-06-2026

Abstract: This study developed and evaluated socio-cultural digital teaching materials for elementary mathematics and examined their validity, practicality, and effectiveness in improving students' problem solving and self-efficacy. We employed a Research & Development design (Borg & Gall; 7 modified stages). A one-group pretest–posttest field try-out was conducted with Grade-5 students (n = 45). Instruments included a Polya-based problem-solving test and a self-efficacy questionnaire (5-point Likert); expert validation covered content, media, and language. Analyses comprised descriptive statistics, Shapiro–Wilk, paired t-test, N-Gain, effect size (Cohen's dz), reliability (Cronbach's α) and inter-rater agreement (Cohen's κ), and Pearson correlation. Findings indicate very high validity (mean 91.1%) and high practicality (teachers 90.2%; students 88.0%; on-task observation 86.4%). Problem-solving performance rose from 55.3 (SD 11.2) to 76.8 (SD 10.1) with N-Gain = 0.48 (moderate), t(44) = 8.75; p < .001; dz = 1.30; the largest gains appeared in understand and look back. Self-efficacy increased from 3.10 (SD 0.52) to 3.78 (SD 0.49) with t(44) = 7.36; p < .001; dz = 1.10. Changes in self-efficacy correlated moderately and positively with problem-solving N-Gain (r = 0.41; p = 0.005). The results support the combined use of local cultural contexts, multimedia-learning principles, and Polya-based strategy scaffolding as a feasible and impactful approach to strengthening mathematical competence and self-belief in elementary students.

Keywords: Multimedia learning, problem solving, self-efficacy, socio-cultural

Abstrak: Penelitian ini bertujuan mengembangkan dan menguji bahan ajar digital berbasis socio-cultural untuk matematika sekolah dasar serta menilai validitas, kepraktisan, dan efektivitasnya dalam meningkatkan pemecahan masalah dan efikasi diri siswa. Desain yang digunakan adalah *Research & Development* model Borg & Gall (modifikasi 7 tahap). Uji lapangan menggunakan desain *one-group pretest–posttest* pada siswa kelas V (n = 45; semester genap 2024/2025). Instrumen meliputi tes pemecahan masalah berbasis indikator Polya dan angket efikasi diri (Likert 1–5); validasi ahli dilakukan pada aspek konten, media, dan bahasa. Analisis mencakup statistik deskriptif, Shapiro–Wilk, paired t-test, N-Gain, effect size (Cohen's dz), reliabilitas (Cronbach's α) dan inter-rater agreement (Cohen's κ), serta korelasi Pearson. Hasil menunjukkan validitas sangat tinggi (rerata 91,1%) dan kepraktisan tinggi (respon guru 90,2%; siswa 88,0%; observasi on-task 86,4%). Instrumen andal (α_pre = 0,82; α_post = 0,85; κ = 0,82). Kemampuan pemecahan masalah meningkat dari 55,3 (SD 11,2) menjadi 76,8 (SD 10,1) dengan N-Gain = 0,48 (sedang), t(44) = 8,75; p < 0,001; dz = 1,30; kenaikan terbesar muncul pada tahap *understand* dan *look back*. Efikasi diri naik dari 3,10 (SD 0,52) menjadi 3,78 (SD 0,49) dengan t(44) = 7,36; p < 0,001; dz = 1,10. Perubahan efikasi diri berasosiasi sedang-positif dengan N-Gain pemecahan masalah (r = 0,41; p = 0,005). Temuan menegaskan bahwa integrasi konteks budaya lokal dengan prinsip multimedia learning dan strategy *scaffolding* Polya merupakan pendekatan yang layak dan berdampak untuk penguatan kompetensi dan keyakinan diri matematis di SD.

Kata kunci: Pembelajaran multimedia, pemecahan masalah, efikasi diri, sosial-budaya

*Corresponding author: nurlaela.2421038@students.um.ac.id

INTRODUCTION

Indonesia's numeracy achievement in PISA 2022 remains substantially below the OECD average, with the official results released on December 5, 2023, and summarized in the updated 2024 country factsheet. Numeracy, as a core component of mathematical competence, encompasses the ability to apply mathematical concepts, procedures, and reasoning to solve problems in real-life contexts, making it inseparable from mathematical modelling, reasoning, and problem-solving skills. The PISA results indicate that the proportion of Indonesian students reaching Levels 5–6 (top performers) in mathematics is nearly zero, signaling persistent weaknesses in higher-order thinking and non-routine problem solving that should be addressed from the elementary school level (OECD, 2023). Strengthening students' early exposure to meaningful mathematical problem-solving tasks is therefore crucial, as elementary education serves as the foundational stage for developing both cognitive competence and positive dispositions toward mathematics. These data collectively highlight an urgent instructional need: to improve instructional practices that not only promote procedural proficiency, but also cultivate conceptual and strategic reasoning through meaningful and contextual tasks. In this sense, mathematics education in Indonesia requires transformative learning designs that integrate real-world situations and stimulate students' cognitive engagement from the earliest stages of schooling.

In response to these challenges, the design of digital teaching materials has gained increasing attention, particularly as meta-analyses in educational technology consistently show that multimedia and digital tools can enhance learning outcomes when they are aligned with established cognitive principles, such as segmentation, signaling, and cognitive load management, and when they provide opportunities for interactive practice and feedback (Dita et al., 2023; Mustofa et al., 2024; Hillmayr et al., 2020). Empirical studies on digital learning environments, such as the Wayang Outpost intelligent tutoring system, demonstrate that adaptive feedback and multimodal representations can support students' mathematical performance by facilitating conceptual understanding and strategy use (Arroyo et al., 2014). Importantly, such systems are not presented as the focus of this study but rather as illustrative evidence that well-designed digital environments, when grounded in learning theory, can meaningfully support mathematics learning. These findings underscore the potential of digital instruction to guide student thinking, support metacognitive processes, and accommodate individual learning differences, especially when digital resources are structured intentionally to reduce cognitive barriers. Taken together, previous research indicates that effective digital instruction does not merely digitize traditional worksheets, but instead transforms the nature of learning activity through intentional design grounded in psychological and pedagogical principles.

Beyond cognitive outcomes, affective factors play a critical role in students' engagement and success in mathematics. Mathematics self-efficacy, defined as students' beliefs in their capability to successfully perform mathematical tasks, has been shown to be consistently and positively associated with achievement, persistence, and strategy selection (Ayotola & Adedeji, 2009; Ding et al., 2024). Large-scale analyses of NAEP data across Grades 4–12 further indicate that variations in self-efficacy are closely related to achievement differences and performance gaps between student groups (Yang et al., 2024). Moreover, formative feedback that is perceived as useful by students can strengthen self-efficacy, thereby indirectly supporting learning outcomes through enhanced motivation and self-regulation (Rakoczy et al., 2019). These findings suggest that effective mathematics instruction should simultaneously address cognitive processes and affective dimensions. In other words, interventions designed solely to improve procedural fluency may fail to produce sustained gains unless they also cultivate students' confidence, autonomy, and willingness to engage productively with mathematical challenge.

Within the Indonesian context, the need to integrate cognitive and affective considerations is further intensified by cultural and linguistic diversity. Culturally responsive mathematics teaching (CRMT) emphasizes the use of meaningful, contextually relevant tasks that draw on students' local practices, experiences, and cultural artifacts. A synthesis of K–12 CRMT research indicates that such approaches increase participation, engagement, and access to learning opportunities, particularly for students from diverse language and cultural backgrounds (Abdulrahim & Orosco, 2020; Song & Coppersmith, 2020). For elementary mathematics, embedding socio-cultural contexts into problem-solving tasks can enhance task relevance and support students in constructing problem representations that are both mathematically and socially meaningful. The integration of cultural context can also offer affective benefits by validating learners' identities and experiences, potentially strengthening both self-efficacy and perseverance. Yet, despite this theoretical support, digital teaching materials developed within Indonesian elementary mathematics classrooms have rarely embedded socio-cultural design principles as a core component of instruction.

Despite these advances, several gaps remain evident in the existing literature. First, many elementary teaching material development studies using Research and Development (R&D) approaches tend to stop at evaluating validity and practicality, without simultaneously examining effectiveness in terms of Polya's problem-solving process indicators (understand, devise a plan, carry out, and look back) and students' mathematical self-efficacy using standardized pretest–posttest designs and effect size measures. Furthermore, meta-analyses of word problem-solving (WPS) interventions largely emphasize cognitive performance indicators, such as accuracy or test scores, while affective constructs like self-efficacy receive comparatively limited attention (Kong et al., 2021; Myers et al., 2022). Second, research has only recently provided a comprehensive mapping of task characteristics that influence WPS difficulty, including lexical inconsistency, unknown placement at the beginning of problems, the need for realistic considerations, and multiple operational steps.

Although these task characteristics are highly relevant for instructional design, their direct translation into the development of elementary digital teaching materials remains limited in R&D-focused studies (Vessonen et al., 2024). As a result, many instructional products do not yet systematically align task design with known sources of problem-solving difficulty faced by elementary learners. Third, studies that integrate multimedia learning principles, socio-cultural task design, and explicit strategy scaffolding (such as Polya's heuristics or schema-based instruction) within a single digital teaching material package for elementary mathematics are still rare. Equally limited are quantitative analyses that examine the relationship between changes in students' self-efficacy (ΔSE) and their learning gains, such as normalized gain (N-Gain), despite growing evidence of the central role of metacognition in mathematics achievement and the effectiveness of metacognition-based interventions (Hillmayr et al., 2020; Sercenia & Prudente, 2023; Xie et al., 2024). This gap suggests the need for integrative research that connects instructional design, cognitive outcomes, and affective development.

Based on these considerations, the present study develops socio-cultural digital teaching materials for elementary mathematics that are aligned with multimedia learning principles to reduce extraneous cognitive load and that incorporate explicit strategy scaffolding based on Polya's problem-solving framework. The study aims to examine the validity and practicality of the developed materials through expert validation, classroom observation, and teacher–student responses; to evaluate their effectiveness in improving students' mathematics problem-solving ability, both in total scores and across Polya's stages, using a one-group pretest–posttest field try-out with effect size and N-Gain analyses; to assess the impact of the intervention on students' mathematical self-efficacy and the reliability of the self-efficacy scale; and to analyze the relationship between changes in self-efficacy (ΔSE) and gains in problem-solving performance (N-Gain) in order to map the contribution of affective factors to cognitive achievement. Through this integrated approach, the study seeks to address both instructional and psychological dimensions of elementary mathematics learning in the Indonesian context. This work therefore offers a novel contribution by synthesizing socio-cultural design elements, multimedia learning theory, Polya-based scaffolding, and self-efficacy analysis within a single instructional product and research model, offering new insight into how digital mathematics learning environments can simultaneously support cognitive growth and affective development.

METHOD

This study adopted the Borg & Gall Research & Development (R&D) model, modified into seven stages: needs analysis, product design, initial prototype development, expert validation, limited try-out, product revision, and field try-out (Borg & Gall, 2003). The R&D approach was chosen to produce socio-cultural digital teaching materials and to test the product's validity, practicality, and effectiveness within an elementary school context. The research was conducted at SDN 023 Tana Tidung during the second semester of the 2024/2025 academic year. The subjects were fifth-grade students selected through purposive sampling: 15 students for the limited try-out and 45 students for the field try-out. The selection criteria were the students' access to digital devices and the relevance of the cultural context.

The development of the socio-cultural digital teaching materials followed a Research and Development (R&D) procedure adapted from Borg and Gall and implemented through seven sequential stages. The process began with a comprehensive needs analysis that examined curriculum demands, characteristics of elementary students, local socio-cultural contexts, and common misconceptions encountered in mathematics problem solving. This stage was intended to ensure that the developed materials were aligned with curriculum standards while remaining meaningful and accessible to students' lived experiences. Based on the needs analysis, the product design stage was conducted by developing a detailed content blueprint, designing contextual word-problem tasks, determining formative and summative assessment components, and integrating multimedia learning principles to manage cognitive load effectively (Mayer, 2009). The next stage involved prototype development, during which the teaching materials were produced in digital form using multimodal representations, such as text, visuals, and guided prompts, to support essential cognitive processing. The prototype was then subjected to expert validation involving specialists in content, media, and language to evaluate accuracy, clarity, and usability. After revisions based on expert feedback, a limited try-out was conducted to examine usability, task flow, and initial classroom feasibility. Further revisions were made to refine the product before implementing a field try-out using a one-group pretest–posttest design aimed at evaluating the effectiveness of the materials in improving students' problem-solving ability and mathematical self-efficacy.

To evaluate the quality and effectiveness of the developed teaching materials, several instruments were employed in this study. Students' mathematical problem-solving ability was measured using a test constructed according to Polya's four problem-solving stages—understanding the problem, devising a plan, carrying out the plan, and looking back—each assessed through an analytic scoring rubric to capture performance at different process levels (Polya, 1945). Mathematical self-efficacy was measured using a questionnaire based on Bandura's social-cognitive theory, encompassing four sources of self-efficacy: mastery experiences, vicarious experiences, social persuasion, and physiological or affective states. The questionnaire used a five-point Likert scale ranging from strongly disagree to strongly agree (Bandura, 1997). In addition, expert validation sheets were used to assess content accuracy, media quality, and language clarity of the teaching materials, while practicality observation sheets were employed to document classroom implementation, task flow, and student engagement. Validation was conducted by three experts, and the reliability analysis of the self-efficacy questionnaire yielded a Cronbach's Alpha coefficient of $\alpha = 0.82$, indicating good internal consistency and meeting the acceptable reliability threshold of ≥ 0.70 (Nunnally & Bernstein, 1994).

The data collection procedure was carried out systematically during the field try-out phase. Prior to the implementation of the digital teaching materials, students were administered a pretest to measure their initial problem-solving ability and mathematical self-efficacy. The intervention was then implemented in regular classroom instruction, with the teacher delivering the lessons using the developed materials and receiving technical assistance from the researcher when necessary. This approach was intended to maintain instructional authenticity while ensuring proper use of the digital product. Following the completion of the instructional sequence, students were given a posttest consisting of the same problem-solving test and self-efficacy questionnaire to measure changes in both cognitive and affective outcomes. In addition to test data, observations were conducted throughout the implementation to assess practicality aspects, such as ease of use, student participation, and task flow. Feedback from both teachers and students was also collected to provide qualitative support for the practicality evaluation.

Data analysis was conducted to evaluate the validity, practicality, and effectiveness of the developed teaching materials. Validity and practicality data obtained from expert judgments, observations, and user responses were analyzed using descriptive statistics in the form of means and percentages, which were then interpreted according to established feasibility categories. To assess learning outcome effectiveness, prerequisite analyses were first conducted, including a Shapiro–Wilk normality test on the difference scores between pretest and posttest results. Improvement in students' problem-solving performance was calculated using normalized gain (N-Gain), computed as $g = (\text{post} - \text{pre}) / (100 - \text{pre})$, and categorized according to established criteria (Hake, 1998). Statistical significance of pretest–posttest differences was examined using a paired t-test for normally distributed data of total scores and for each Polya problem-solving stage. Practical significance was assessed using effect size measures, namely Cohen's d for paired t-tests (Cohen, 1988). For self-efficacy data, descriptive analyses were conducted for each dimension, reliability was re-examined using Cronbach's Alpha, and pretest–posttest comparisons with effect size reporting were performed.

The results of validity, practicality, and effectiveness analyses were interpreted based on predetermined feasibility categories, N-Gain criteria, statistical significance, and effect size magnitude. The outcomes of these analyses are presented in the Results and Discussion section to show how the needs analysis, product design, expert validation, practicality evaluation, and field try-out findings were connected systematically.

RESULTS AND DISCUSSION

The results of this study are presented according to the sequential stages of the development process, beginning with the needs analysis and product design, followed by expert validation, practicality evaluation, and field try-out effectiveness. This structure is intended to clarify how the initial empirical findings informed the instructional design of the socio-cultural digital teaching materials and how the product was subsequently evaluated in classroom implementation.

Table 1. Summary of needs analysis findings

Aspect	**Key Finding**	**Implication for Product Design**
Curriculum demands	Grade 5 mathematics learning outcomes under the Merdeka Curriculum require students not only to compute correctly but also to reason through the complete word-problem-solving process—understanding, planning, executing, and checking	Teaching materials needed to explicitly scaffold Polya's four problem-solving stages rather than emphasize procedural drills alone
Characteristics of elementary students	Fifth-grade students rely heavily on concrete, visual, and contextually familiar cues to interpret the abstract relationships expressed in word problems	Tasks were designed with concrete visual representations, everyday language, and step-by-step guided prompts
Local socio-cultural context	Students' daily lives in Tana Tidung are closely connected to local livelihoods such as river transportation, small-scale trading, and agricultural activities, providing authentic situations that could be translated into mathematical tasks	Contextual word problems were built around locally familiar scenarios (e.g., river travel time, market transactions, harvest calculations) instead of generic contexts
Common problem-solving misconceptions	Students frequently had difficulty identifying relevant information within a problem, selecting an appropriate solution strategy, and checking whether their final answer was reasonable	Materials incorporated guided prompts and reflective "look back" questions that specifically targeted these three recurring difficulties

To make the outcomes of this stage explicit, Table 1 summarizes the key findings obtained from the needs analysis across the four aspects examined during the development procedure—curriculum demands, characteristics of elementary students, local socio-cultural context, and common problem-solving misconceptions—together with their corresponding implications for the design of the socio-cultural digital teaching materials.

As presented in Table 1, the needs analysis identified four interrelated areas—curriculum demands, student characteristics, local socio-cultural context, and common misconceptions—that together informed the design of the socio-cultural digital teaching materials. These findings became the empirical basis for combining contextual tasks, visual representations, guided prompts, and reflective questions in the developed product, ensuring that the material was aligned with curriculum standards while remaining meaningful and accessible to students' lived experiences.

Based on the needs analysis, the product design was organized into a digital teaching material package integrating three main components: socio-cultural contexts, multimedia learning principles, and Polya-based strategy scaffolding. The socio-cultural component was reflected in contextual word problems adapted from students' local environment and everyday experiences. The multimedia component was developed through the use of concise text, visual representations, signaling, and segmented learning activities to reduce unnecessary cognitive load. Meanwhile, the strategy scaffolding component guided students to analyze the problem, plan a solution, perform mathematical procedures, and evaluate their answers systematically. The design also included formative tasks, reflection prompts, and feedback-oriented activities to support both problem-solving ability and mathematical self-efficacy.

These results indicate that the product design was not developed merely by digitizing conventional worksheets, but by transforming instructional content into a structured learning environment grounded in students' needs, cultural relevance, and cognitive support. The needs analysis and design process therefore served as an important bridge between the theoretical framework and the subsequent validation and field try-out stages. In this sense, the socio-cultural digital teaching materials were designed to address both cognitive barriers in mathematical problem solving and affective aspects related to students' confidence in learning mathematics.

Following the needs analysis and product design stages, the feasibility of the socio-cultural digital teaching materials was evaluated in terms of validity and practicality prior to their implementation in the field try-out phase. Expert judgment focused on three main aspects: content accuracy, media feasibility, and language clarity. In addition, practicality indicators were examined through teacher and student responses as well as classroom observations, with emphasis on ease of use, task flow, and student engagement during learning activities. The purpose of this analysis was to determine whether the product met the required academic, pedagogical, and technical standards for classroom implementation. These results are summarized in Table 2, which presents the average expert ratings for each aspect and the corresponding feasibility categories.

Table 2. Validity and practicality product

Aspect	Average (%)	Category
Content	92.4	Very valid
Media	90.1	Very valid
Language	90.7	Very valid
Total	91.1	Very valid

The findings presented in Table 2 indicate that the developed materials achieved a "very valid" classification, with the total average rating amounting to 91.1%. This outcome demonstrates that the product meets academic and technical requirements, especially in terms of curricular alignment, clarity of language, and visual design compatibility with learning objectives. The consistently high scores across content, media, and language categories serve as a strong indicator that the material was coherent and pedagogically appropriate. These results further suggest that the design features embedded in the product—such as multimodal representations and socio-cultural contextualization—were considered relevant, meaningful, and functional by expert reviewers. In addition, the practicality data captured through teacher and student responses showed that the product could be implemented smoothly and effectively in the classroom environment. Taken together, these findings confirmed that the digital teaching materials were ready to be tested in real instructional settings, strengthening the rationale for moving forward into the effectiveness analysis.

Before conducting analyses on student learning outcomes, the reliability of the measurement instruments was examined to ensure that data collected during the study were accurate and internally consistent. This step was essential because unreliable measurement tools can significantly reduce the validity of research findings and lead to misleading conclusions. Reliability testing focused on two main instruments: the self-efficacy questionnaire, which measured student beliefs about their mathematical competence, and the problem-solving rubric, used to assess performance based on Polya's indicators. Cronbach's α was applied to examine internal consistency of the questionnaire at both pretest and posttest stages, while Cohen's κ was used to evaluate inter-rater agreement on the rubric. Table 3 presents the reliability scores associated with each instrument, confirming that all values exceeded commonly accepted threshold levels. These results indicated that both instruments were stable and consistent across measurement points, reducing the risk of error in subsequent data interpretation. The high reliability further implies that observed changes in the results were more likely attributable to learning intervention effects rather than measurement inconsistencies.

As shown in Table 3, the self-efficacy scale demonstrated Cronbach's α values of 0.82 for the pretest and 0.85 for the posttest, suggesting good to excellent internal consistency. The increase in alpha value from pretest to posttest indicates stronger coherence in student responses following the intervention, possibly reflecting improved understanding or more stable perception of their mathematical abilities. Similarly, the Cohen's κ value of 0.82 showed that the problem-solving rubric yielded strong inter-rater agreement. This finding is particularly important because rubric scoring often involves subjective interpretation, and high κ values indicate minimal discrepancies between evaluators. These reliability results provide compelling evidence that both instruments were functioning effectively as tools for measuring student performance and belief patterns. Accordingly, the reliability outcomes serve as a methodological assurance that the instruments generated dependable data for subsequent statistical procedures.

Table 3. Instrument reliability

Instrument	Construct Measured	Scale/ Score	Time	N	Reliability Index	Value	Category/ Interpretation
Self-efficacy Questionnaire	Mathematical self-confidence	Likert 1–5	Pretest	45	Cronbach's α	0.82	Reliable (good)
Self-efficacy Questionnaire	Mathematical self-confidence	Likert 1–5	Posttest	45	Cronbach's α	0.85	Reliable (very good)
Problem-solving Rubric	Problem-solving performance (based on Polya)	Score 0–100	—	—	Cohen's κ	0.82	Very good agreement

After confirming reliability, the study proceeded to examine the effectiveness of the teaching materials in improving students' problem-solving performance using the one-group pretest–posttest design. The analysis focused on Polya's four indicators: understand, devise a plan, carry out, and look back. Each indicator was analyzed using paired t-tests, effect sizes, and normalized gain (N-Gain) values. Prior to inferential analysis, normality assumptions for the difference scores were tested and confirmed, supporting the validity of the statistical procedures. Table 4 displays the descriptive and inferential statistics associated with pretest–posttest comparisons across all four indicators. The substantial increases in mean scores and the large effect sizes provide convincing evidence that the intervention meaningfully enhanced students' mathematical problem-solving skills. These improvements reflect not only increased accuracy but also stronger strategic reasoning and metacognitive engagement.

Table 4. Comparison of pretest–posttest scores for problem-solving ability

Polya Indicator	Pre (M±SD)	Post (M±SD)	Δ	N-Gain	$t(44)$	p	dz
Understand	58.4±14.2	80.3±12.6	21.9	0.53	8.01	<0.001	1.25
Devise a plan	52.7±13.8	74.1±12.9	21.4	0.45	7.54	<0.001	1.15
Carry out	50.1±12.7	70.4±12.1	20.3	0.41	6.92	<0.001	1.03
Look back	48.3±13.5	72.5±12.8	24.2	0.47	7.86	<0.001	1.18

As shown in Table 4, all four indicators demonstrated significant improvement, with score increases ranging from 20.3 to 24.2 points. The highest gain was observed in the "look back" stage, indicating that students developed stronger capabilities in reviewing and evaluating their solutions. Effect size values above 1.00 suggest that improvements were not only statistically significant but also pedagogically meaningful. Additionally, medium-range N-Gain values further reinforce that the intervention was effective in promoting learning progress across process phases. These findings align with prior evidence indicating that structured strategy instruction can enhance mathematical reasoning and metacognitive awareness. The consistent pattern of improvement across indicators demonstrates that the intervention supported students not merely in executing procedures, but in engaging more deeply with problem structure and reflection.

To provide a broader perspective on learning gains, normalized gain values were further classified into low, medium, and high categories. This approach offers a more nuanced interpretation of student progress by examining performance variability rather than relying solely on averages. Table 5 presents the distribution of students across N-Gain categories. These findings are important because they show not only overall effectiveness but also how improvements were distributed across individuals.

Table 5. N-gain distributor category

Category	**Range**	**Total (n)**	**Percentage**
Low	$g < 0.30$	9	20.0%
Medium	$0.30 \leq g \leq 0.70$	28	62.2%
High	$g > 0.70$	8	17.8%
Total		45	100%

Table 5 reveals that most students (62.2%) achieved medium N-Gain scores, while 17.8% reached high-level improvement. Meanwhile, 20.0% of students were classified in the low category, indicating that further instructional differentiation may be beneficial for some individuals. Taken together, these data strengthen the interpretation that the intervention produced substantial learning gains for the majority of participants, but also highlight opportunities for targeted support to maximize outcomes.

In addition to cognitive performance, this study also investigated learning effects on mathematical self-efficacy, which reflects students' beliefs in their ability to succeed in mathematical tasks. Changes in self-efficacy were measured across four dimensions—mastery experience, vicarious experience, social persuasion, and physiological/affective states—allowing a detailed analysis of affective development. Table 6 summarizes the pretest–posttest differences across these dimensions. Improvements in each domain suggest that the intervention influenced not only what students could do, but also how they perceived themselves as learners.

Table 6. Improvement in mathematical self-efficacy

Dimension	**Pre (M±SD)**	**Post (M±SD)**	**Δ**
Mastery experience	3.05±0.64	3.88±0.57	+0.83
Vicarious experience	3.10±0.60	3.70±0.55	+0.60
Social persuasion	3.20±0.58	3.82±0.52	+0.62
Physiological/affective states	3.09±0.59	3.57±0.53	+0.48
Total	3.10±0.52	3.78±0.49	+0.68

As shown in Table 6, the largest increase occurred in mastery experience, which is consistent with Bandura's theory that successful task performance strengthens confidence. Gains were also evident in vicarious experience and social persuasion, indicating that collaborative learning and teacher feedback contributed positively to self-belief. Improvements in physiological/affective responses suggest that students experienced reduced anxiety or increased comfort when engaging in mathematical tasks. These findings reinforce the idea that digital learning environments can support affective growth in addition to cognitive development. To examine alignment with the social-cognitive framework, the relationship between improvements in self-efficacy and problem-solving gains was analyzed using Pearson correlation. Table 7 presents this association, demonstrating how affective and cognitive developments interacted within the learning process.

Table 7. Relationship between self-efficacy and learning gains

Variable	**n**	**Test**	**Coefficient (r)**	**95% CI for r**	***p***
ΔSelf-Efficacy (ΔSE) vs N-Gain *problem solving*	45	Pearson correlation	0.41	[0.13; 0.63]	0.005

As shown in Table 7, a medium-positive and statistically significant correlation ($r = 0.41$; $p = 0.005$) was found between changes in self-efficacy and N-Gain. This finding indicates that students who experienced larger increases in self-confidence tended to achieve stronger learning gains in problem solving. Although the correlation does not establish causality, the result aligns with theoretical expectations that self-efficacy influences persistence, strategy selection, and self-regulation. This pattern highlights the integrative role of affective variables in shaping cognitive outcomes.

The high average expert ratings in the “very valid” category, together with strong practicality indicators, indicate that the developed socio-cultural digital teaching materials were appropriate for classroom implementation and aligned with the learning needs identified during the preliminary stage. The needs analysis showed that students required learning materials that could help them interpret contextual word problems, organize relevant information, select suitable strategies, and evaluate the reasonableness of their answers. Therefore, the validation results should not be interpreted only as evidence of product feasibility, but also as confirmation that the content, language, media features, and instructional flow were consistent with the intended problem-solving support. Within the multimedia learning framework, clear signaling, well-structured multimodal representations, and cognitive load management supported students in processing mathematical information more effectively (Hillmayr et al., 2020; Mayer, 2009). In addition, the integration of socio-cultural contexts made the tasks more meaningful and accessible for students, particularly because the mathematical problems were connected to familiar situations and local experiences (Abdulrahim & Orosco, 2020; Song & Coppersmith, 2020). This finding suggests that the product design addressed both the cognitive needs of students in solving mathematical problems and the contextual relevance required for meaningful learning. Thus, the result is consistent with meta-analytic evidence showing that technology-based interventions are more effective when they are theoretically grounded and connected to meaningful learning tasks (Hillmayr et al., 2020).

The significant improvement in total problem-solving scores, supported by a large effect size, demonstrates that the intervention enhanced students’ mathematical problem-solving performance beyond procedural calculation. The N-Gain value in the medium category indicates that most students experienced meaningful progress during the field try-out, although the improvement varied across individuals. This result corresponds with previous findings on word-problem-solving interventions at the elementary level, which show that structured instructional support can improve students' ability to solve mathematical word problems (Kong et al., 2021; Myers et al., 2022), a pattern consistent with broader meta-analytic evidence that problem-based learning approaches produce positive effects on learners' cognitive and skill outcomes (Dochy et al., 2003). In this study, the improvement can be explained by the product's design, which guided students through contextual tasks, visual representations, and Polya-based prompts. These features helped students move from simply attempting to calculate an answer toward understanding the problem situation, planning a solution, implementing the procedure, and checking the result. Therefore, the observed increase in problem-solving ability reflects the instructional function of the developed product in supporting more systematic mathematical reasoning.

A more detailed analysis of Polya's indicators showed that the strongest gains occurred in the "understand" and "look back" stages. This pattern is important because the needs analysis indicated that students often struggled to identify relevant information and rarely reviewed their answers after solving a problem. Through the digital teaching materials, students were repeatedly guided to recognize known and unknown information, connect the problem with visual or socio-cultural representations, and reflect on whether their answers were reasonable. This finding suggests that the intervention strengthened students' metacognitive engagement, particularly in representing the problem and evaluating the solution. Such a pattern is commonly found in interventions that emphasize heuristic guidance and metacognitive scaffolding (Chan et al., 2023; Xie et al., 2024). The improvement in these stages also indicates that students were supported in avoiding premature closure, namely the tendency to stop after obtaining an answer without checking its mathematical meaning or accuracy (Schoenfeld, 1985). The coherence of these results is further supported by evidence that schema-based instruction and structured strategy practice can improve elementary students' word-problem performance (Fuchs et al., 2021; Powell et al., 2015). In digital learning environments, multimodal representations and targeted feedback can further facilitate students' strategy selection and execution (Arroyo et al., 2014).

In the affective domain, the significant increase in self-efficacy scores accompanied by a large effect size indicates that the product also supported students' confidence in learning mathematics. This improvement can be understood through the learning experiences embedded in the intervention. Contextual tasks that were familiar to students allowed them to experience gradual success in solving problems, thereby strengthening mastery experience. Collaborative activities enabled students to observe peers' strategies, while teacher feedback provided social persuasion that encouraged students to believe in their own mathematical ability. These mechanisms are consistent with Bandura's view that self-efficacy develops through mastery experiences, vicarious experiences, social persuasion, and physiological or affective states (Bandura, 1997). Prior studies also show that self-efficacy is positively associated with mathematics achievement and can be strengthened through instructional approaches that combine cognitive strategy support and formative feedback (Ayotola & Adedeji, 2009; Rakoczy et al., 2019). Furthermore, when multimedia design principles reduce unnecessary cognitive load, students may experience mathematics tasks as more manageable, which in turn can support both learning effectiveness and self-efficacy development (Mayer, 2009).

Taken together, the medium-positive association between changes in self-efficacy (ΔSE) and problem-solving learning gains (N-Gain) suggests that students' affective and cognitive development occurred in an interconnected manner during the intervention. Students who demonstrated greater increases in self-confidence tended to achieve stronger gains in problem-solving performance. Although this correlation does not establish causality, it provides empirical support for the assumption that self-belief is related to students' persistence, strategy selection, and self-regulation when solving challenging mathematical problems (Ding et al., 2024; Richardson et al., 2012). In the context of this study, the developed product may have contributed to this relationship through two complementary pathways. First, Polya-based scaffolding, visual representations, and contextual tasks helped students process mathematical problems more systematically. Second, successful engagement with these tasks strengthened students' belief that they were capable of solving mathematics problems. This finding reinforces the importance of designing mathematics learning materials that simultaneously support cognitive strategy use and affective readiness.

From a practical standpoint, the findings suggest that socio-cultural digital teaching materials can be used by teachers to strengthen mathematical problem solving when the design explicitly connects local contexts, visual representations, and structured strategy prompts. Contextualized problem-solving tasks grounded in students' local practices or cultural artifacts can help students understand problem situations more easily, especially when combined with worked examples and reflective prompts that strengthen the "understand" and "look back" processes (Chan et al., 2023; Myers et al., 2022). Teachers are also encouraged to provide specific, strategy-focused formative feedback rather than simple "right or wrong" judgments, because such feedback can support both learning interest and self-efficacy (Rakoczy et al., 2019). In addition, multimedia principles such as signaling and segmentation should be applied to reduce extraneous cognitive load and support deeper processing (Hillmayr et al., 2020; Mayer, 2009). Cultural responsiveness, achieved by embedding local contexts and role models as sources of vicarious experience, can further strengthen the relevance and durability of students' learning experiences (Abdulrahim & Orosco, 2020; Purwati et al., 2019; Song & Coppersmith, 2020).

Finally, the interpretation of these findings should consider the methodological limitation of the study. Because the field try-out used a one-group pretest–posttest design without a comparison group, the observed improvements cannot be attributed solely to the intervention with full causal certainty. Future studies are therefore recommended to employ quasi-experimental designs, such as non-equivalent control groups or stepped-wedge models, and to include delayed post-tests to examine retention and transfer effects. Further investigation into the mediating role of self-efficacy and the moderating influence of implementation fidelity and device access would also contribute to a more comprehensive understanding of how digital, socio-culturally grounded teaching materials influence mathematics learning outcomes (Hillmayr et al., 2020; Powell et al., 2015).

CONCLUSION

This study concludes that the socio-cultural digital teaching materials developed in this research are feasible for elementary mathematics learning and have the potential to support students' mathematical problem-solving ability and self-efficacy. The development process showed that integrating local socio-cultural contexts into digital teaching materials can make mathematical problems more meaningful, familiar, and accessible for elementary students. The validation and practicality findings indicate that the product was pedagogically appropriate, technically usable, and relevant to classroom learning needs.

The implementation of the product in the field try-out suggests that socio-cultural digital teaching materials can support students in engaging more systematically with mathematical problem solving. Through contextual tasks, multimodal representations, and Polya-based scaffolding, students were encouraged to understand problem situations, plan solution strategies, carry out procedures, and reflect on the reasonableness of their answers. The findings also indicate that the product did not merely support procedural performance, but also strengthened students' strategic reasoning and metacognitive awareness during problem solving.

In addition to cognitive development, the study shows that the designed learning environment contributed to students' mathematical self-efficacy. Contextual problems, guided activities, peer interaction, and teacher feedback provided learning experiences that helped students feel more capable of completing mathematical tasks. This suggests that students' confidence in mathematics can be strengthened when learning materials provide structured support, meaningful contexts, and opportunities for successful problem-solving experiences. Overall, this research highlights the value of combining socio-cultural context, multimedia learning principles, and strategy scaffolding in elementary mathematics instruction. The main contribution of the study lies in demonstrating that Research and Development studies should not stop at product validity and practicality, but can be extended to examine how instructional products influence both cognitive and affective learning outcomes. In this study, problem-solving development and self-efficacy growth appeared to be mutually connected, suggesting that effective mathematics learning design should address both students' thinking processes and their beliefs about their own capabilities.

Nevertheless, the interpretation of the findings should consider the limitation of the one-group pretest–posttest design. Future studies are recommended to examine the product using comparison groups, delayed post-tests, and more advanced analyses to explore the possible mediating role of self-efficacy and the influence of implementation fidelity or access to digital devices. Practically, the findings suggest that teachers can strengthen mathematics learning by designing contextual digital materials that include clear visual guidance, structured problem-solving prompts, and formative feedback that supports both competence and self-belief.

REFERENCES

Abdulrahim, N. A., & Orosco, M. J. (2020). Culturally responsive mathematics teaching: A research synthesis. *Urban Review*, *52*, 1–25. 10.1007/s11256-019-00509-2

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, *24*(4), 387-426. https://doi.org/10.1007/s40593-014-0023-y

Ayotola, A., & Adedeji, T. (2009). The relationship between mathematics self-efficacy and achievement in mathematics. *Procedia – Social and Behavioral Sciences*, *1*(1), 953–

957. 10.1016/j.sbspro.2009.01.172
Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.
Borg, W. R., & Gall, M. D. (2003). *Educational research: An introduction* (7th ed.). Allyn & Bacon.
Chan, S.-W., Looi, C.-K., Ho, W. K., & Kim, M. S. (2023). Tools and Approaches for Integrating Computational Thinking and Mathematics: A Scoping Review of Current Empirical Studies. *Journal of Educational Computing Research*, *60*(8), 2036–2080. https://doi.org/10.1177/07356331221098793
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). LawrenceErlbaum Associates.
Ding, Y., Klapp, A., & Yang Hansen, K. (2024). The importance of mathematics self-concept and self-efficacy for mathematics achievement: A comparison between public and independent schools in Sweden. *Educational Psychology*, *44*(8), 872-892. 10.1080/01443410.2024.2410217
Dita, K. I., Tuririday, H. T., Damopolii, I., & Latjompoh, M. (2023). Designing the human circulatory system e-module to increase student achievement. *Inornatus: Biology Education Journal*, *3*(2), 75–84. https://doi.org/10.30862/inornatus.v3i2.422
Dochy, F., Segers, M., Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, *13*, 533–568. 10.1016/S0959475202000257
Fuchs, L. S., Seethaler, P. M., Sterba, S. K., Craddock, C., Fuchs, D., Compton, D. L., Geary, D. C., & Changas, P. (2021). Closing the Word-Problem Achievement Gap in First Grade: Schema-Based Word-Problem Intervention with Embedded Language Comprehension Instruction. *Journal of Educational Psychology*, *113*(1), 86–103. https://doi.org/10.1037/edu0000467
Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809
Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S., & Reiss, K. (2020). The potential of digital tools to enhance mathematics and science learning. *Computers & Education*, *153*, 103897. 10.1016/j.compedu.2020.103897
Kong, J. E., Yan, C., Serceki, A., & Swanson, H. L. (2021). Word-Problem-Solving Interventions for Elementary Students With Learning Disabilities: A Selective Meta-Analysis of the Literature. *Learning Disability Quarterly*, *44*(4), 248–260. https://doi.org/10.1177/0731948721994843
Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
Mustofa, A., Hayuana, W., Damopolii, I., Ibrohim, I., & Susilo, H. (2024). The discovery learning and Google sites: Its application in learning the process of urine formation for high school students. *Inornatus: Biology Education Journal*, *4*(2), 132–150. https://doi.org/10.30862/inornatus.v4i2.711
Myers, J. A., Witzel, B. S., Powell, S. R., Li, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). A meta-analysis of mathematics word-problem solving interventions for elementary students who evidence mathematics difficulties. *Review of Educational Research*, *92*(5), 695–742. https://doi.org/10.3102/00346543211070049
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

OECD. (2023). *PISA 2022 Results*. OECD.
Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
Powell, S. R., Fuchs, L. S., Cirino, P. T., Fuchs, D., Compton, D. L., & Changas, P. (2015). Effects of a multitier support system on calculation, word-problem, and prealgebraic performance among at-risk learners. *Exceptional Children*, *81*(4), 443–470. https://doi.org/10.1177/0014402914563702
Purwati, P., Marasabessy, F., & Damopolii, I. (2019). Enhancing students activity and problem solving skill through CTL-based local wisdom approach. *Journal of Physics: Conference Series*, *1321*(3), 032077. https://doi.org/10.1088/1742-6596/1321/3/032077
Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*, *60*, 154–165. https://doi.org/10.1016/j.learninstruc.2018.01.004
Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838
Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.
Sercenia, J. C., & Prudente, M. S. (2023). Effectiveness of the Metacognitive-based Pedagogical Intervention on Mathematics Achievement: A Meta-Analysis. *International Journal of Instruction*, *16*(4), 561–578. https://doi.org/10.29333/iji.2023.16432a
Song, K. H., & Coppersmith, S. (2020). Working toward linguistically and culturally responsive math teaching through a year-long urban teacher training program for English learners. *Journal of Urban Mathematics Education*, *13*(1), 1–28. https://doi.org/10.5281/zenodo.4145481
Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Korhonen, J., Aunio, P., & Laine, A. (2024). Task Characteristics Associated with Mathematical Word Problem-Solving Performance Among Elementary School-Aged Children: A Systematic Review and Meta-Analysis. *Educational Psychology Review*, *36*(4), 117. https://doi.org/10.1007/s10648-024-09954-2
Xie, Y., Cao, J., Qi, C., Li, X., & Zhao, Y. (2024). A meta-analysis of the relationship between metacognition and mathematics achievement. *Acta Psychologica*, *241*, 104032. 10.1016/j.actpsy.2024.104486
Yang, Y., Maeda, Y., & Gentry, M. (2024). The relationship between mathematics self-efficacy and mathematics achievement: Multilevel analysis with NAEP 2019. *Large-Scale Assessments in Education*, *12*(1), 16. https://doi.org/10.1186/s40536-024-00204-z