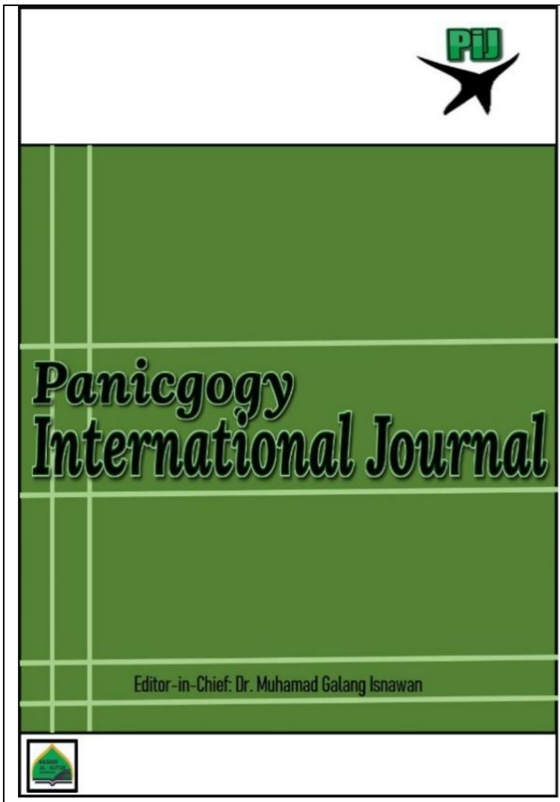
Panicgogy International Journal

Publication details, including instructions for authors and subscription information:
https://nakiscience.com/index.php/pij



Designing and validating a CTL-based social arithmetic module to foster junior high school students' creative thinking and intrinsic motivation

Mugiono Mugiono[1], Yumiati Yumiati[2], Sendi Ramdhani[3]

[1]Department of Mathematics Education, Universitas Terbuka, Jakarta, Indonesia, mugi16ono@gmail.com

[2]Department of Mathematics Education, Universitas Terbuka, Jakarta, Indonesia, yumi@ecamput.ut.ac.id

[3]Department of Primary Education, Universitas Terbuka, Jakarta, Indonesia, sendi@ecampus.ut.ac.id

To cite this article:
Mugiono, M., Yumiati, Y., & Ramdhani, S. (2026). Designing and validating a CTL-based social arithmetic module to foster junior high school students' creative thinking and intrinsic motivation. *Panicgogy International Journal, 4*(2), 110 – 132.
To link to this article:
https://nakiscience.com/index.php/pij/
Published by:
Nashir Al-Kutub Indonesia
Residential Street Kila Rengganis, Block I, Number 11, Labuapi, Indonesia, 83361

Panicgogy International Journal, 4(2), 110 – 132, 2026
https://nakiscience.com/index.php/pij

Original Article

Designing and validating a CTL-based social arithmetic module to foster junior high school students' creative thinking and intrinsic motivation

Mugiono Mugiono[1*], Yumiati Yumiati[2], Sendi Ramdhani[3]
[1]Department of Mathematics Education, Universitas Terbuka, Jakarta, Indonesia, mugi16ono@gmail.com
[2]Department of Mathematics Education, Universitas Terbuka, Jakarta, Indonesia, yumi@ecamput.ut.ac.id
[3]Department of Primary Education, Universitas Terbuka, Jakarta, Indonesia, sendi@ecampus.ut.ac.id

*Corresponding Author: mugi16ono@gmail.com

Abstract
Creative thinking and intrinsic motivation remain insufficiently developed in Indonesian junior high school mathematics, as reflected in students' low performance in the PISA 2022 mathematics and creative-thinking domains. This development research, adapted from Plomp's model, comprised preliminary research, prototyping, and assessment phases. It aimed to design, validate, and evaluate a Contextual Teaching and Learning (CTL)-based Social Arithmetic module for fostering seventh-grade students' creative thinking and intrinsic motivation. Formative evaluation included a one-to-one evaluation with three students, a small-group trial with six students, and a field test involving 28 heterogeneous seventh-grade students at SMP Avava Batam. Data were collected using expert validation sheets, classroom observations, student and teacher response questionnaires, a pretest–posttest creative-thinking test, and an intrinsic-motivation questionnaire. Aiken's V was used to assess validity, while practicality and learning improvements were analyzed descriptively and through normalized gain. The module was valid, achieving mean validity scores of 71.11% for media, 77.08% for content, and 72.92% for language. It was also practical to very practical, receiving positive responses of 80.51% from students and 70.00% from the teacher. Students' creative-thinking scores increased, particularly in originality, while intrinsic motivation achieved a medium N-gain of 0.46. However, because the field test used a one-group pretest–posttest design without inferential analysis, the results indicate descriptive improvement rather than causality. Further studies should confirm these findings through controlled designs, inferential testing, and larger, more diverse samples across schools and regions.
Keywords: Contextual Teaching and Learning; Teaching Module; Creative Thinking; Intrinsic Motivation; Social Arithmetic

1. Introduction

Creative thinking is widely regarded as one of the higher-order cognitive competencies that twenty-first-century mathematics education must deliberately cultivate, since it underlies students' capacity to generate multiple ideas, shift between strategies, and produce original and well-elaborated solutions to non-routine problems (Hendriana & Fadhillah, 2019; Suherman & Vidákovich, 2022b). Within the broader framework of twenty-first-century competencies, creative thinking is positioned alongside critical thinking, communication, and collaboration as one of the "4Cs" that determine students' readiness for future work and civic life (Thornhill-Miller et al., 2023). In mathematics specifically, creative thinking is most visibly expressed when students face open-ended, contextual tasks that admit more than one correct solution path, because such tasks compel them to exercise fluency, flexibility, originality, and elaboration rather than merely reproduce memorized procedures (Rahayuningsih et al., 2021; Gunawan et al., 2022; Anggareni & Hidayat, 2022).

Empirical evidence, however, suggests that Indonesian junior high school students' creative thinking remains underdeveloped. The Programme for International Student Assessment (PISA) 2022, which for the first time assessed creative thinking as an innovative domain alongside mathematics, reported that Indonesian fifteen-year-olds obtained a mathematics score of 366, far below the average of countries in the Organization for Economic Co-operation and Development, together with an equally modest performance in the creative-thinking domain (OECD, 2023). These figures corroborate classroom-level observations that students routinely struggle when a mathematics task departs from a single, procedural solution path, particularly in topics such as Social Arithmetic that are conceptually simple but pedagogically demanding because they require translating everyday transactional situations into flexible mathematical reasoning.

Alongside creative thinking, intrinsic motivation is a second competency of central concern, because it is the internal driving force that sustains students' engagement with challenging, open mathematical tasks over time (Ryan & Deci, 2020). Self-Determination Theory posits that intrinsic motivation flourishes when learning activities are experienced as interesting, autonomy-supportive, and personally meaningful rather than externally imposed (Ryan & Deci, 2020; Knopik & Oszwa, 2023). Empirical studies confirm that students' self-determined motivation is positively associated with mathematics achievement and inversely associated with mathematics anxiety (Iyamuremye et al., 2023), and that instructional designs that foreground authentic inquiry, such as the 5E framework combined with inquiry-based learning, can measurably strengthen both motivation and engagement in the mathematics classroom (Al Tulyani et al., 2026). Creative thinking and intrinsic motivation are therefore best understood not as isolated outcomes but as plausibly mutually reinforcing: a task that invites exploration and multiple solution strategies is also a task that is more likely to be experienced as interesting and self-relevant, and vice versa — a proposition this study is positioned to illustrate descriptively, though not to test directly (see Section 4).

A substantial body of prior research has examined the Contextual Teaching and Learning (CTL) approach and its close relative, Realistic Mathematics Education (RME), as vehicles for connecting mathematics instruction to students' everyday experiences. Studies have shown that CTL-based comics improve critical thinking and character (Lestari et al., 2021), that CTL-based e-modules on sequences and series raise students' learning interest (Martin et al., 2021), and that RME-based worksheets and teaching tools strengthen students' mathematical reasoning, communication, and higher-order thinking across various topics (Erita et al., 2022; Sa'id et al., 2021; Febriyani et al., 2024; Purwitaningrum & Prahmana, 2021; Saleh et al., 2018). A recent meta-analysis further confirms that RME approaches consistently improve a range of cognitive mathematical outcomes across the last two decades of research (Juandi et al., 2022). Within Social Arithmetic specifically, RME-based modules have likewise been shown to enhance students' mathematical literacy (Malinda et al., 2025) and to connect naturally with financial-literacy competencies through project-based learning (Sagita et al., 2022).

Closer inspection of this literature, however, reveals three specific limitations that the present study is positioned to address. First, almost all of the CTL/RME development studies cited above report only a single quality outcome — either a cognitive outcome such as reasoning or critical thinking, or an affective outcome such as interest — rather than jointly documenting a cognitive and an affective competency produced by the same validated instructional product. Second, where an affective outcome is examined, it is rarely intrinsic motivation defined and measured through Self-Determination Theory's autonomy-competence- relevance framework (Ryan & Deci, 2020; Knopik & Oszwa, 2023); studies more commonly rely on adjacent but conceptually broader constructs such as general interest or engagement (Al Tulyani et al., 2026). Third, none of the Social-Arithmetic-specific studies identified above report a full development-research triangulation of validity, practicality, and effectiveness — formal content-validity indices from multiple expert judges, triangulated practicality data from both students and teachers, and a quantified learning gain — within a single systematically validated module. Examining creative thinking and intrinsic motivation jointly, using this triangulated evidentiary structure, therefore constitutes a substantive rather than incremental contribution: it tests directly whether a single contextual design can move a cognitive and an affective outcome together, a question the cited literature has left open by design rather than by oversight. This gap is pedagogically consequential because Social Arithmetic sits at the intersection of procedural computation and everyday financial reasoning, making it a natural but underexploited context for cultivating both flexible mathematical thinking and self-determined motivation.

To address this gap, the present study designs, develops, and empirically validates a CTL-based teaching module on Social Arithmetic using the development research model proposed by Plomp (2013), which structures the design process into preliminary research, prototyping, and assessment phases and has been shown to be well suited to producing instructional materials whose validity, practicality, and effectiveness can be documented systematically (Alamanda et al., 2023). The module deliberately embeds authentic buying-and-selling scenarios, open-ended profit-maximization tasks, and reflective questioning to operationalize the CTL principles of constructivism, inquiry, and authentic assessment, with the explicit aim of eliciting fluency, flexibility, originality, and elaboration in students' problem solving while simultaneously nurturing their sense of autonomy, competence, and relevance in the mathematics classroom.

Guided by this rationale, the study addresses two research questions: (1) How is a Contextual Teaching and Learning-based module on Social Arithmetic designed and developed through the Plomp model to support seventh-grade students' creative thinking and intrinsic motivation? (2) To what extent is the resulting module valid, practical, and effective in enhancing seventh-grade students' creative thinking skills and intrinsic motivation? Because this second question integrates three distinct product-quality outcomes, each is treated as a separate sub-question mapped to its own data source, criterion, and analysis: (2a) *validity*, assessed from expert validation sheets and analyzed with Aiken's V; (2b) *practicality*, assessed from observation sheets and student/teacher response questionnaires and analyzed descriptively; and (2c) *effectiveness*, assessed descriptively from the pretest-posttest creative-thinking test and intrinsic-motivation questionnaire and further quantified with the normalized gain (N-gain) index. Section 3 reports the results following this same 2a–2b–2c structure.

2. Methods

2.1 Research Design

This study employed a development research (design research) approach adapted from the Plomp model, chosen because the primary goal was to produce a validated instructional product — a CTL-based teaching module on Social Arithmetic — rather than merely to test a hypothesis about an existing treatment (Plomp, 2013; Alamanda et al., 2023). The Plomp model organizes the design process into three iterative phases: preliminary research, in which curriculum documents, the target concept, and students' characteristics are analyzed to identify the instructional problem; the prototyping phase, in which successive prototypes of the module are produced and refined through a layered formative evaluation consisting of self-evaluation, expert review, one-to-one evaluation, and small-group evaluation; and the assessment phase, in which the refined prototype is field-tested to determine its practicality and effectiveness under authentic classroom conditions. Table 1 links each phase to its formative-evaluation activity, participant group, data source, and the revision decision it produced, so that the process summarized narratively in Section 3.1 can also be traced in a single reproducible reference table.

Table 1

Plomp Development Phases, Formative-Evaluation Activities, Participants, Data Sources, and Revision Decisions

Phase	Formative-Evaluation Activity	Participants	Data Source	Revision Decision
Preliminary research	Curriculum and needs analysis; diagnostic test	26 seventh-grade students; classroom teacher (interview)	Diagnostic creative-thinking test; teacher interview	Identified a low creative-thinking baseline and a procedure-only teaching pattern; informed the design of Prototype I around open-ended profit-maximization tasks
Prototyping	Self-evaluation and expert review	Researcher; three expert validators	Expert validation sheets (media, content, language) with qualitative comments	Redesigned cover; added contextual examples and digital-media illustrations; varied font styles for formulas → Prototype II
Prototyping	One-to-one evaluation	Three students (low, medium, high ability), class VII A	Structured observation; student comments on instructions	Clarified ambiguous instructions and terminology → refinement toward Prototype III
Prototyping	Small-group evaluation	Six students of varying ability, class VII B	Practicality questionnaire (ease of use, feasibility, appeal, time)	Refined pacing and activity clarity → Prototype III (field-test version)
Assessment	Field test	28 students, class VII C; 1 teacher; 2 observers	Observation sheets; student/teacher response questionnaires; creative-thinking pretest-posttest; intrinsic-motivation questionnaire	Confirmed validity; generated the practicality and effectiveness data reported in Section 3

This design was considered appropriate for the study's dual concern with creative thinking and intrinsic motivation because it allows validity, practicality, and effectiveness to be examined as three distinct but related quality criteria rather than as a single global judgment. Validity was established through structured expert judgment prior to classroom use; practicality was triangulated from observation sheets and questionnaire responses collected from both students and the teacher during actual instruction; and effectiveness was established through a pretest-posttest comparison of students' creative-thinking scores and intrinsic-motivation questionnaire scores following instruction with the module. Each phase produced concrete revision decisions that were carried forward into the next prototype, so that the final field-tested version reflects a cumulative process of evidence-based refinement rather than a single, untested design.

2.2 Participants

The study was conducted at SMP Avava Batam, and its subjects were seventh-grade students distributed across the three formative-evaluation stages required by the Plomp model. The one-to-one evaluation involved three students purposively selected from class VII A to represent low-, medium-, and high-ability learners, so that idiosyncratic comprehension problems in the module's instructions and language could be detected early. The small-group trial involved six students of varying ability from class VII B, selected to examine the module's practicality — ease of use, instructional feasibility, appeal, and time efficiency — under conditions that more closely approximate normal classroom use than the one-to-one stage.

The field test, which generated the effectiveness and practicality data reported in this article, involved one intact and heterogeneous class of twenty-eight students from class VII C, taught by their regular mathematics teacher following a lesson design aligned with the module's CTL components; two independent observers monitored implementation across all meetings. In addition to the student subjects, three validators — two mathematics education lecturers and one practicing junior high school mathematics teacher — served as expert judges during the validation stage, providing scored judgments and qualitative comments on the module's media, content, and language aspects prior to classroom trial.

Participants at every stage were selected using purposive sampling. Eligibility required current enrollment in the seventh grade at SMP Avava Batam and prior completion of the prerequisite arithmetic material; for the one-to-one and small-group stages, students were additionally selected on teacher-nominated ability classification (low, medium, high) so that the module was trialed across the full ability range before the field test. Class VII C was selected for the field test because its regular teacher volunteered to implement the CTL-based lesson design and because the class's general ability distribution was judged by the school to be representative of the school's seventh-grade classes as a whole. The three validators were purposively selected for relevant expertise: two held a graduate qualification in mathematics education with prior experience validating instructional materials, and one was an in-service junior high school mathematics teacher with at least five years of experience teaching Social Arithmetic.

2.3 Ethical Clearance

This study received ethical clearance from Universitas Terbuka prior to data collection. Because all student participants were minors, written informed consent was obtained from parents or guardians together with students' own assent to participate, and informed consent was separately obtained from the classroom teacher; all participants were informed of their right to withdraw at any time without consequence. Student data were anonymized prior to analysis, and only aggregated results are reported in this article.

2.4 Data Collection

Four categories of instruments were used to collect data across the development stages. Validity data were obtained from expert validation sheets covering the media, content, and language aspects of the module, each rated on a four-point scale from strongly disagree to strongly agree. Practicality data were obtained from three complementary sources: an observation sheet completed by two observers during the field test to record the implementation of each CTL component, a Guttman-scaled (yes/no) student response questionnaire administered after the field test, and an equivalent teacher response questionnaire completed after the teacher had used the module across the instructional sequence.

Effectiveness data on creative thinking were obtained from an essay-based test consisting of two contextual Social Arithmetic problems, each scored against three indicators — fluency, flexibility, and novelty (originality) — using a rubric developed specifically for this study. Each indicator was scored on a 0–4 scale (0 = no relevant response; 4 = a complete, well-elaborated, and mathematically correct response satisfying the indicator), yielding a maximum raw score of 12 per item and 24 across the two items. The first item asked students to propose as many valid combinations of purchases and payment methods as possible within a given spending constraint (fluency), to solve the problem using at least two different strategies (flexibility), and to propose one original strategy — not demonstrated in the module — for maximizing profit (novelty); the second item posed a structurally different but comparably open discount-and-profit scenario.

2.5 Data Analysis

Validity data were analyzed using Aiken's V content-validity coefficient, in which each validator's rating is converted into a validity index for every item and then averaged across items and validators to yield an overall validity index for the module, following procedures that have been widely used and further refined for confirming instrument and product content validity in mathematics education research (Merino-Soto, 2023; Kania et al., 2024). Following Merino-Soto's (2023) reporting convention, the resulting Aiken's V index for each aspect was rescaled to a 0–100% metric (V × 100) to aid interpretation against the four-category classification used in this study; the percentages reported in Table 3 therefore represent rescaled Aiken's V values — computed from the ratings of three validators on the four-point scale described in Section 2.3 — rather than raw mean scores. Item-level V values are available from the corresponding author on request. The resulting percentage was classified as very valid (86–100%), valid (76–85%), fairly valid (60–75%), or less valid (below 60%). Practicality data from the observation sheets and the two response questionnaires were analyzed descriptively as mean percentages of the maximum possible score and classified into five categories ranging from impractical to very practical.

Effectiveness data on creative thinking were analyzed descriptively by comparing students' mean pretest and posttest percentage scores against a five-level classification ranging from very low to very high. Effectiveness data on intrinsic motivation were analyzed in two complementary ways: first, students were classified into low, medium, and high motivation categories using a norm-referenced criterion (mean ± 1 SD) established once from the pretest distribution and held constant when classifying posttest scores, so that any category shift reflects genuine change in students' raw scores rather than a re-scaled distribution at posttest; second, the magnitude of change from pretest to posttest was quantified using the normalized gain (N-gain) index, computed as the ratio of the actual score increase to the maximum possible increase, a measure that remains one of the most widely used and defensible indicators of instructional effectiveness in pretest-posttest designs (Coletta & Steinert, 2020). N-gain values were classified as high ($\geq$ 0.70), medium (0.30–0.69), or low ($<$ 0.30).

Because the field test employed a one-group pretest-posttest design without a comparison group or a formal inferential test, the pretest-posttest comparisons reported in Sections 3.3 and 3.4 are interpreted as a descriptively observed improvement rather than a statistically confirmed causal effect. Accordingly, the term "effective" is used throughout this article in this descriptive sense — a positive, educationally meaningful change consistent with the module's N-gain classification — and not to claim formal statistical significance; future replications are encouraged to complement this descriptive analysis with an appropriate inferential test (e.g., a paired-samples t-test or, where normality is not met, the Wilcoxon signed-rank test) and a standardized effect size.

3. Results

3.1 Development of the CTL-Based Module

The preliminary research stage identified a clear instructional problem: a curriculum and needs analysis, together with a diagnostic test given to 26 seventh-grade students, showed a low average creative-thinking score, and teacher interviews confirmed that Social Arithmetic was routinely taught through worked examples followed by procedurally similar exercises, leaving little room for open exploration. This analysis directly informed the design of Prototype I, a CTL-based module structured around authentic buying-and-selling scenarios and an open-ended profit-maximization task requiring students to propose multiple purchase combinations, alternative profit strategies, and one genuinely original idea for increasing profit in the long term.

During the prototyping phase, self-evaluation and expert review generated concrete revisions, including an increase in the number of contextual examples derived from a single real-life problem, the addition of daily-life illustrations and digital media references, a redesigned cover with visuals relevant to Social Arithmetic, more extensive use of illustrative images in the concept-explanation section, and varied font styles to highlight key formulas, producing Prototype II. The one-to-one evaluation, conducted with three students of differing ability, further refined ambiguous instructions and terminology that had proven difficult to interpret, while the small-group evaluation with six students refined the pacing and clarity of activities related to ease of use, feasibility, appeal, and time allocation, yielding Prototype III. This prototype was the version implemented in the field test with twenty-eight students, whose results are reported below. Table 2 consolidates this narrative into an explicit revision log linking each identified issue to its source of feedback, the modification made, and the resulting prototype; representative pages of the module illustrating these revisions — the redesigned cover and the addition of contextual illustrations in the practice exercises — are presented in Figure 1 and Figure 2.

Table 2

Revision Log Across Prototyping Stages

Issue Identified	Source of Feedback	Modification Made	Resulting Prototype
Cover and visual design perceived as unengaging	Expert review (media aspect)	Redesigned cover with Social-Arithmetic-relevant visuals; added illustrative images throughout the concept-explanation section	Prototype II
Limited number of contextual examples anchored to a single scenario	Self-evaluation; expert review (content aspect)	Increased the number of contextual examples derived from a single real-life problem; added digital-media references	Prototype II
Key formulas not visually distinguishable from surrounding text	Expert review (language/media aspect)	Applied varied font styles to highlight key formulas	Prototype II

Issue Identified	Source of Feedback	Modification Made	Resulting Prototype
Ambiguous task instructions and unfamiliar terminology	One-to-one evaluation (three students)	Rewrote instructions in simpler language; replaced unfamiliar terms with everyday equivalents	Prototype III (draft)
Uneven pacing; some activities perceived as too time-consuming	Small-group evaluation (six students)	Adjusted activity sequencing and time allocation; simplified lower-priority tasks	Prototype III (field-tested version)

Figure 1
Redesigned Cover of the CTL-Based Module Following Self-Evaluation and Expert Review (Prototype I vs. Prototype II)

(a) Prototype I — initial cover, before revision

(b) Prototype II — redesigned cover with Social-Arithmetic-relevant illustration

As shown in Figure 1, the initial cover (Prototype I) used a generic layout with no visual reference to the module's content, whereas the redesigned cover (Prototype II) incorporates illustrations directly tied to Social Arithmetic — a shopping table, a discount tag, receipts, and a rising bar chart — giving students an immediate visual cue to the module's real-world, transaction-based focus. This revision followed the media-aspect feedback summarized in Table 2.

Figure 2

Addition of Contextual Illustrations in the Practice Exercises (Uji Pemahaman), Before and After Revision

(a) Before revision — only the first item is accompanied by a contextual photograph

(b) After revision — a second contextual photograph was added to illustrate the construction-materials item

Figure 2 illustrates the revision made in response to feedback that not every practice item was accompanied by a supporting image: a photograph relevant to the second item (a contractor purchasing cement) was added, consistent with the increase in contextual, image-supported examples reported in Table 2.

Taken together, this iterative sequence directly answers the study's first research question: the module was designed and refined through a Plomp-based cycle of needs analysis, expert-informed revision, and learner-based tryout before its quality was formally assessed.

3.2 Validity of the Module

To answer the study's first research question regarding the quality of the developed module, three validators with expertise in mathematics education evaluated Prototype I across three aspects: media, content, and language. Each aspect was rated against a set of validation indicators and analyzed using Aiken's V content-validity coefficient, which was then rescaled to a 0–100% metric for ease of interpretation. Table 3 summarizes the resulting validity indices for each aspect together with the overall mean validity score.

Table 3

Summary of Expert Validation Results (Rescaled Aiken's V)

Aspect	Mean Score	Maximum Score	Percentage	Criterion
Media	42.66	60	71.11%	Valid
Content	49.33	64	77.08%	Valid
Language	35.00	48	72.92%	Valid
Overall Mean	-	-	73.70%	Valid

As shown in Table 3, all three aspects of the module were judged valid by the three expert validators, with percentages ranging narrowly between 71.11% and 77.08%. The content aspect obtained the highest rating, suggesting that the validators were particularly convinced by the alignment between the Social Arithmetic tasks and the learning objectives, while the media aspect, though still valid, received comparatively more critical feedback concerning visual design. A closer inspection of the three validators' individual scores also revealed moderate variability, with one validator rating several items more conservatively than the other two; because the overall validity index is computed as an aggregate across all validators rather than from any single judge's score, this variability did not compromise the module's overall valid classification, but it did generate specific, item-level revision notes that were incorporated before the field test (see Table 2).

3.3 Effectiveness of the Module in Enhancing Creative Thinking

Students' creative thinking was assessed with a two-item, indicator-based test administered as a pretest before, and a posttest after, the field-test implementation of the module, covering the fluency, flexibility, and novelty (originality) indicators for each item. Raw scores were converted into percentages and disaggregated by indicator to allow a fine-grained comparison of pretest and posttest performance. Table 4 presents these pretest–posttest percentages together with the resulting gain, in percentage points, for each indicator and item.

Table 4

Pretest-Posttest Results of Students' Creative Thinking by Indicator

Item No.	Indicator	Pretest (%)	Posttest (%)	Gain (pp)
1	Fluency	54.76	84.52	+29.76
	Flexibility	50.00	83.33	+33.33
	Novelty (Originality)	46.43	84.52	+38.09
2	Fluency	60.71	89.29	+28.58
	Flexibility	61.90	85.71	+23.81
	Novelty (Originality)	57.14	82.14	+25.00
Overall (weighted by raw score)	All indicators	50.00	85.00	+35.00
Criterion	-	Low	High	-

Note. The overall pretest and posttest percentages (50.00% and 85.00%) are computed as the ratio of students' total raw scores across all indicators and both items to the total maximum possible raw score, rather than as the unweighted arithmetic mean of the six indicator percentages shown above; because Item 1 and Item 2 differ slightly in their per-indicator maximum raw scores, this weighted overall percentage does not necessarily equal the simple average of the six displayed values (55.16% for the pretest and 84.92% for the posttest). Both computations are reported here for transparency; the exact per-student raw-score matrix underlying the weighted figures is available from the corresponding author on request.

Table 4 shows that students' creative-thinking scores rose substantially on every indicator across both test items. The largest gains occurred in the novelty (originality) indicator, which increased by 38.09 percentage points on the first item and 25.00 percentage points on the second, exceeding the gains observed for fluency and flexibility on the same items. Because originality is generally the indicator most resistant to short-term instructional improvement, this pattern provides descriptive evidence that the module's open, non-routine tasks — particularly the profit-maximization item that explicitly asked students to propose an unconventional strategy — succeeded in creating genuine opportunities for original mathematical thinking rather than merely rehearsing fluency in generating answers. These results are consistent with the effectiveness component of the study's second research question with respect to creative thinking: the module was associated with a descriptive improvement in creative thinking, most pronounced for the most cognitively demanding indicator of creativity; a formal inferential test is recommended in future replications to confirm statistical significance (see Section 2.4).

3.4 Effectiveness of the Module in Enhancing Intrinsic Motivation

Students’ intrinsic motivation toward mathematics learning was measured with a validated questionnaire administered as a pretest before, and a posttest after, the field-test implementation of the module. Raw scores were classified into low, medium, and high categories using cut-off points derived from the pretest mean and standard deviation, which were then held constant when classifying the posttest scores. Table 5 displays the resulting distribution of students across these three motivation categories at both testing points.

Table 5

Distribution of Students’ Intrinsic Motivation Categories

Category	Pretest n (%)	Posttest n (%)
Low	15 (53.57%)	0 (0.00%)
Medium	9 (32.14%)	13 (46.43%)
High	4 (14.29%)	15 (53.57%)
Total (N)	28	28

The low, medium, and high motivation cut-off scores reported in Table 5 were established once, from the pretest distribution (mean ± 1 SD), and held constant when classifying both pretest and posttest scores, so that the category shift reported below reflects genuine change in students’ raw scores rather than a re-scaled distribution at posttest. The pretest mean was 61.82 and the posttest mean was 73.39.

Figure 3

Distribution of Students’ Intrinsic Motivation Categories Before and After Using the CTL-Based Module (N = 28)

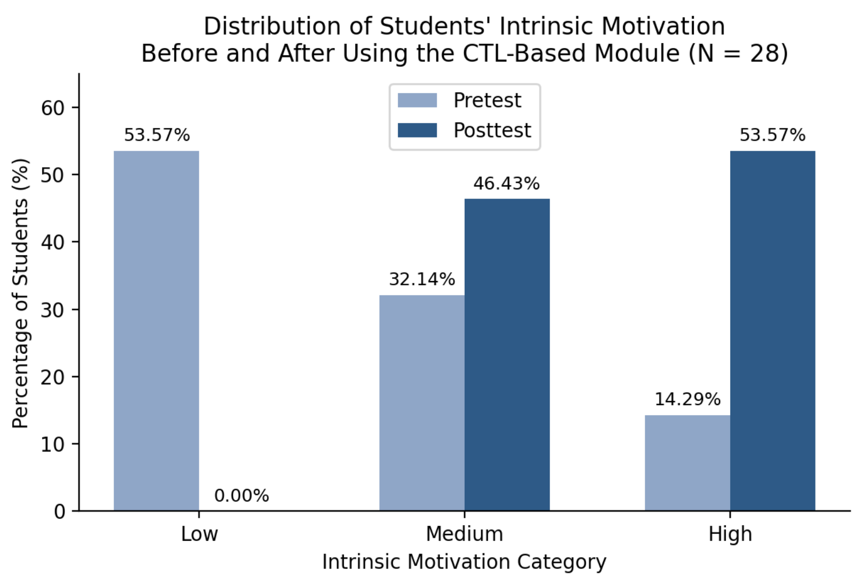


Table 5 and Figure 3 show a marked shift in the distribution of students' intrinsic motivation: the proportion of students in the low category fell from 53.57% to 0.00%, while the proportion in the high category rose from 14.29% to 53.57%. This pattern indicates that no student remained in the low-motivation category after instruction with the module, and that more than half of the class moved into the high-motivation category. As presented in Table 6, this shift corresponds to an increase in the mean motivation score from 61.82 in the pretest to 73.39 in the posttest, yielding a normalized gain of 0.46, classified as medium. The medium, rather than high, N-gain is consistent with the substantial proportion of students who began in the low category and therefore needed the largest change to reach the higher bands; nonetheless, the complete elimination of the low category is itself a practically meaningful outcome. Together, these findings are consistent with the motivation component of the study's second research question: the module was associated with a descriptive improvement in students' intrinsic motivation, with a medium normalized gain.

Table 6

Results of the Normalized Gain (N-Gain) Test for Intrinsic Motivation

Test	N	Mean	Mean Difference	N-Gain	Criterion
Pretest	28	61.82	11.57	0.46	Medium
Posttest	28	73.39			

3.5 Practicality of the Module

The practicality of the module was evaluated from two complementary sources of data: response questionnaires completed by students during the small-group trial and the field test, and a response questionnaire completed by the class teacher during the field test. Each respondent group rated the module across several aspects, including ease of use, implementation, appeal, and time efficiency, with responses analyzed descriptively as percentages classified against a standard practicality criterion. Table 7 summarizes these percentage scores and their corresponding practicality classifications for each respondent group and aspect.

Table 7

Summary of Practicality Results from the Small-Group Trial and the Field Test

Respondent / Aspect	Percentage	Criterion
Small group – ease of use	92.18%	Very practical
Small group – implementation	77.50%	Practical
Small group – appeal	76.25%	Practical
Small group – time efficiency	77.23%	Practical
Small group – overall	78.92%	Practical
Field test – student overall	80.51%	Very practical
Teacher – ease of use	66.67%	Practical

Respondent / Aspect	Percentage	Criterion
Teacher – implementation	73.33%	Practical
Teacher – appeal	66.67%	Practical
Teacher – time efficiency	73.33%	Practical
Teacher – overall	70.00%	Practical

Table 7 shows that the module was rated practical to very practical across all sources of practicality data. Students in the small-group trial rated the module 78.92% overall, with the ease-of-use aspect reaching 92.18% (very practical), while students in the field test rated the module even more favorably overall, at 80.51% (very practical). The teacher's rating was more conservative at 70.00% (practical), with implementation and time efficiency rated somewhat higher than ease of use and appeal, suggesting that the additional facilitation demands of orchestrating an open, contextual lesson were more salient to the teacher than to the students. Considered alongside the validity and effectiveness results reported above, these findings jointly answer the study's second research question: the CTL-based Social Arithmetic module is valid and practical, and was associated with a descriptive improvement in seventh-grade students' creative thinking and intrinsic motivation.

4. Discussion

The finding that the CTL-based Social Arithmetic module achieved a valid classification across its media, content, and language aspects (71.11%, 77.08%, and 72.92%, respectively) corroborates a substantial body of development research indicating that instructional materials designed through a systematic model such as Plomp's typically attain valid-to-very-valid classifications once expert feedback has been incorporated into successive prototypes (Alamanda et al., 2023). The moderate variability observed among the three validators — one of whom rated aspects of the module more conservatively than the others — illustrates a well-documented feature of expert-judgment validity: individual experts differ in their internal calibration even when they broadly agree on a product's adequacy, which is precisely why aggregated indices such as Aiken's V, rather than any single validator's score, are recommended as the basis for validity decisions (Merino-Soto, 2023; Kania et al., 2024). That the content aspect obtained the highest percentage among the three (77.08%) is also consistent with the module's design priority: because its central innovation lies in translating buying-and-selling scenarios into open mathematical tasks, the validators appear to have judged the substantive alignment between content and contextual purpose more favorably than the surface-level media presentation, an observation that is instructive for future revisions focusing on visual and graphic refinement.

The module's association with an improvement in students' creative-thinking scores from a low pretest average to a high posttest average lends descriptive support to the theoretical claim that creative thinking is most reliably developed when students are given open, contextual tasks that resist a single procedural solution (Rahayuningsih et al., 2021; Gunawan et al., 2022). Notably, the largest gain was observed in the novelty (originality) indicator, which rose by 38.09 and 25.00 percentage points on the two test items respectively. This pattern is theoretically significant because originality is consistently reported in the literature as the most resistant component of mathematical creativity to instructional improvement, since it requires students to depart from conventional strategies rather than merely apply them more fluently (Suherman & Vidákovich, 2022b; Anggareni & Hidayat, 2022). The comparatively large originality gain observed here suggests that the module's profit-maximization task — which explicitly asked students to propose an unconventional idea for increasing profit beyond ordinary pricing strategies — succeeded in creating the kind of genuinely open problem space that the assessment literature identifies as necessary, though rarely achieved, for cultivating originality specifically, rather than fluency and flexibility alone. This finding extends prior CTL-based intervention studies, which have more commonly reported gains in critical thinking or general learning outcomes (Lestari et al., 2021; Martin et al., 2021), by illustrating that a carefully constructed contextual task can also move the more elusive originality component of creative thinking; confirming this pattern inferentially, with a larger and independently validated creative-thinking instrument, remains an important next step (see Section 2.3 and the Limitations below).

The complete disappearance of the low-motivation category (from 53.57% to 0.00%) and the more than three-fold increase in the high-motivation category (from 14.29% to 53.57%) after instruction with the module is well explained through the lens of Self-Determination Theory, which holds that intrinsic motivation is nurtured when learners experience autonomy, competence, and relevance in their activities rather than external control (Ryan & Deci, 2020; Knopik & Oszwa, 2023). By grounding Social Arithmetic in transactions that students plausibly encounter outside school — buying clothes, negotiating profit margins — the module gave students a sense that the mathematics they were doing mattered beyond the classroom, a relevance effect that prior studies link directly to stronger self-determined motivation and lower mathematics anxiety (Iyamuremye et al., 2023). The module's open-ended tasks additionally supported a sense of competence by allowing multiple valid pathways to a correct or defensible answer, echoing findings from inquiry-based interventions using the 5E framework, in which structuring lessons around genuine exploration rather than direct transmission measurably increased both motivation and engagement (Al Tulyani et al., 2026). At the same time, the medium N-gain of 0.46, rather than a high gain, is a realistic and theoretically expected outcome: Self- Determination Theory emphasizes that intrinsic motivation is a psychologically embedded orientation that tends to shift gradually rather than abruptly, so a single instructional module implemented across a limited number of meetings would be expected to produce a moderate, not transformative, shift, particularly for the sizeable proportion of students who began in the low-motivation category and needed to move furthest along the continuum.

The module's practicality — rated 78.92% by the small group, 80.51% by the field-test class, and 70.00% by the teacher — situates it within the practical-to-very-practical range reported for comparable CTL- and RME-based instructional materials (Erita et al., 2022; Febriyani et al., 2024; Sa'id et al., 2021). The particularly high ease-of-use rating from students during the small-group stage (92.18%) suggests that the contextual framing of Social Arithmetic problems, being drawn from familiar buying-and-selling situations, lowered the comprehension barrier that abstract algebraic tasks typically pose to seventh-grade learners. The teacher's comparatively more conservative rating (70.00%), by contrast, likely reflects the additional facilitation demands that a CTL-based lesson places on the instructor — managing open-ended discussion, accommodating multiple solution strategies, and orchestrating the seven CTL components within a fixed lesson period — challenges that are well documented in studies of CTL implementation and that point toward a concrete area for future teacher-support materials accompanying the module.

Taken as a whole, these results are consistent with — though they do not directly test — the theoretical proposition, foreshadowed in the introduction, that creative thinking and intrinsic motivation are mutually reinforcing rather than independent outcomes of contextual mathematics instruction. Because the two outcomes were analyzed separately in this study rather than through a joint association analysis, this convergence should be read as a theoretically plausible interpretation of two outcomes that improved together within the same intervention, rather than as evidence of a tested reciprocal relationship. A planned correlational analysis between individual students' creative-thinking gain scores and their motivation gain scores — or a suitable structural model relating the two — is recommended in future replications to establish whether the two competencies are empirically, and not only theoretically, interrelated. With that qualification in place, the same design features that made the module's tasks cognitively open enough to elicit fluency, flexibility, and originality — authentic scenarios, multiple correct pathways, and reflective prompts — are also the features that Self-Determination Theory identifies as conducive to autonomy and relevance, and therefore to intrinsic motivation. This convergence advances the RME and CTL literature, which has more often examined creative thinking (Rahayuningsih et al., 2021; Malinda et al., 2025) and motivation (Al Tulyani et al., 2026; Knopik & Oszwa, 2023) as separate strands of inquiry, by offering descriptive evidence, drawn from a single validated instructional product, that a well-designed contextual module can move both outcomes together. It also responds directly to the twenty-first-century skills literature's call for instructional designs that cultivate creativity as a systemic, assessable competency rather than an incidental byproduct of good teaching

(Thornhill-Miller et al., 2023), by demonstrating one concrete, replicable pathway — embedding realistic financial-transaction contexts into a rigorously validated module — through which that goal can be pursued in an ordinary junior high school classroom.

Several considerations temper these conclusions. The field test involved a single class in one school using a one-group pretest-posttest design without a comparison group or formal inferential test, so the magnitude of the reported gains, particularly the medium N-gain for motivation, should be read descriptively and not be generalized without further replication — ideally incorporating a comparison group, an inferential test such as a paired-samples t-test or Wilcoxon signed-rank test, and a standardized effect size — across schools with different socio-economic and curricular contexts, consistent with cautions raised in broader meta-analytic work on RME interventions (Juandi et al., 2022). Moreover, because the creative-thinking test comprised only two contextual items, the observed indicator-level differences — while theoretically interpretable — warrant confirmation using a longer or independently validated creative-thinking instrument before strong claims are made about which specific CTL component drives originality gains. Nonetheless, the convergence of valid, practical, and descriptively effective classifications obtained through triangulated instruments provides a credible evidentiary basis for recommending the module, and the design principles underlying it, for broader classroom use and further empirical testing.

5. Conclusion

This study has shown that a Contextual Teaching and Learning-based module on Social Arithmetic can be systematically designed and refined through the Plomp development model, progressing from a needs-based preliminary analysis through iterative formative evaluation — self-evaluation, expert review, one-to-one trial, and small-group trial — to a field-tested prototype whose quality is documented rather than assumed. The resulting module was classified as valid on its media, content, and language aspects, with expert-validation percentages of 71.11%, 77.08%, and 72.92% respectively, and was judged practical to very practical by both students (78.92%–80.51%) and the teacher (70.00%) who used it in an authentic classroom setting. These findings directly answer the study's first research question by demonstrating that a rigorous, evidence-based development process can yield a contextual instructional product that is both technically sound and classroom-ready.

With respect to the study's second research question, the module was also associated with a descriptive improvement on both target outcomes: students' creative-thinking scores rose from a low pretest average to a high posttest average, with the largest gains occurring in the originality indicator, and their intrinsic motivation improved from a profile dominated by low motivation to one dominated by high motivation, reflected in a medium normalized gain of 0.46. Because these findings derive from a one-group pretest-posttest design without a comparison group or formal inferential test, they should be treated as promising descriptive evidence rather than a confirmed causal effect pending replication. Beyond these quantitative outcomes, the study contributes theoretically by illustrating that creative thinking and intrinsic motivation can plausibly be cultivated jointly when contextual mathematics tasks are deliberately designed to be open, authentic, and personally relevant — a proposition that future research should test directly through a joint association analysis of the two outcomes. Future research is also encouraged to replicate this design across multiple schools and a longer instructional period with an inferential test and comparison group, to refine teacher-support materials that ease the facilitation demands of CTL-based lessons, and to examine whether the originality gains observed here persist and transfer to Social Arithmetic tasks embedded in different real-life contexts.

Acknowledgments

The authors gratefully acknowledge the principal, mathematics teacher, and seventh-grade students of SMP Avava Batam for their participation and institutional support. The authors also thank the expert validators and classroom observers for their constructive feedback during the development and evaluation of the module.

Declarations

Author Contributions : Mugiono Mugiono: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, and Writing - Original Draft. Yumiati Yumiati: Supervision and Writing - Review & Editing. Sendi Ramdhani: Supervision and Writing - Review & Editing. All authors read and approved the final manuscript.

Funding Statement : This research received no external funding.

Conflict of Interest : The authors declare that they have no conflict of interest.

Ethics Approval : This study received ethical clearance from Universitas Terbuka before data collection. All procedures were conducted in accordance with applicable institutional ethical standards.

Informed Consent : Written informed consent was obtained from the parents or legal guardians of all student participants, together with the students' assent. Informed consent was also obtained from the participating mathematics teacher. Participation was voluntary, and participants could withdraw at any time without consequence.

AI Declaration : During the preparation of this manuscript, the authors used artificial intelligence tools solely for grammar checking and language polishing. All AI-assisted output

was reviewed and edited by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

Data Availability : The data supporting the findings of this study are available from the corresponding author upon reasonable request. Access is subject to appropriate safeguards for participant privacy and confidentiality.

6. References

Al Tulyani, R., Younes, R., & Habre, S. (2026). The impact of using the IBL approach and the 5E framework in mathematics classrooms on students' motivation and engagement. *Teaching Mathematics and its Applications: An International Journal of the IMA*, hrag012. https://doi.org/10.1093/teamat/hrag012

Alamanda, A., Mawardi, M., & Suryani, O. (2023). Development of teaching material based on Plomp development model to support Indonesian Merdeka Curriculum on chemical bonding topic in Phase E. *Jurnal Pijar MIPA, 18*(4), 564–571. https://doi.org/10.29303/jpm.v18i4.5288

Anggareni, P., & Hidayat, A. F. (2022). Students' creative thinking skills on differentiated instruction. *International Journal of Trends in Mathematics Education Research, 5*(4), 365–373. https://doi.org/10.33122/ijtmer.v5i4.161

Coletta, V. P., & Steinert, J. J. (2020). Why normalized gain should continue to be used in analyzing pre-instruction and post-instruction scores on concept inventories. *Physical Review Physics Education Research, 16*(1), Article 010108. https://doi.org/10.1103/PhysRevPhysEducRes.16.010108

Erita, S., Utami, E. S. D., & Ningsih, F. (2022). Realistic mathematics education-based student worksheet to improve students' mathematical reasoning on circle material. *Indonesian Journal of Science and Mathematics Education, 5*(2), 210–223. https://doi.org/10.24042/ijsme.v5i2.11249

Febriyani, A., Habudin, H., & Rachmiati, W. (2024). Development of Realistic Mathematics Education-based learner worksheets (LKPD) to enhance higher-order thinking skills in Grade IV students of SD/MI. *Journal of Research on Mathematics Instruction, 5*(2), 54–63. https://doi.org/10.33578/jrmi.v5i2.95

Gunawan, G., Kartono, K., Wardono, W., & Kharisudin, I. (2022). Analysis of mathematical creative thinking skill: In terms of self-confidence. *International Journal of Instruction, 15*(4), 1011–1034. https://doi.org/10.29333/iji.2022.15454a

Hendriana, H., & Fadhillah, F. M. (2019). The students' mathematical creative thinking ability of junior high school through problem-solving approach. *Infinity Journal, 8*(1), 11–20. https://doi.org/10.22460/infinity.v8i1.p11-20

Iyamuremye, E., Ndayambaje, I., & Muwonge, C. M. (2023). Relationships of mathematics achievement with self-determined motivation and mathematics anxiety among senior two students in Northern Rwanda. *Heliyon, 9*(4), Article e15411. https://doi.org/10.1016/j.heliyon.2023.e15411

Juandi, D., Kusumah, Y. S., & Tamur, M. (2022). A meta-analysis of the last two decades of realistic mathematics education approaches. *International Journal of Instruction, 15*(1), 381–400. https://doi.org/10.29333/iji.2022.15122a

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education), 10*(1), 64–79. https://doi.org/10.21831/reid.v10i1.71032

Knopik, T., & Oszwa, U. (2023). Self-determination in maths education: How to strengthen students' positive attitude to mathematics and develop their emotional/social competences. *Multidisciplinary Journal of School Education, 12*(1), 165–186. https://doi.org/10.35765/mjse.2023.1223.08

Lestari, F. P., Ahmadi, F., & Rochmad, R. (2021). The implementation of mathematics comic through contextual teaching and learning to improve critical thinking ability and character. *European Journal of Educational Research, 10*(1), 497–508. https://doi.org/10.12973/eu-jer.10.1.497

Malinda, A., Suanto, E., & Maimunah, M. (2025). Development of a realistic mathematics education (RME)-based learning module on social arithmetic to enhance students' mathematical literacy. *Jurnal Absis: Jurnal Pendidikan Matematika dan Matematika, 8*(3), 541–555. https://doi.org/10.30606/absis.v8i3.3366

Martin, M., Syamsuri, S., Pujiastuti, H., & Hendrayana, A. (2021). Pengembangan e-modul berbasis pendekatan contextual teaching and learning pada materi barisan dan deret untuk meningkatkan minat belajar siswa SMP. *Jurnal Derivat, 8*(2), 72–87. https://doi.org/10.31316/j.derivat.v8i2.1927

Merino-Soto, C. (2023). Aiken's V coefficient: Differences in content validity judgments. *MHSalud: Revista en Ciencias del Movimiento Humano y Salud, 20*(1), 23–32. https://doi.org/10.15359/mhs.20-1.3

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Plomp, T. (2013). Educational design research: An introduction. In T. Plomp and N. Nieveen (Eds.), *Educational design research* (pp. 10–51). Netherlands Institute for Curriculum Development.

Purwitaningrum, R., & Prahmana, R. C. I. (2021). Developing instructional materials on mathematics logical thinking through the Indonesian realistic mathematics education approach. *International Journal of Education and Learning, 3*(1), 13–19. https://doi.org/10.31763/ijele.v3i1.178

Rahayuningsih, S., Sirajuddin, S., & Ikram, M. (2021). Using open-ended problem-solving tests to identify students' mathematical creative thinking ability. *Participatory Educational Research, 8*(3), 285–299. https://doi.org/10.17275/per.21.66.8.3

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sa'id, I. A., Pambudi, D. S., Hobri, H., Safik, M., & Insani, K. (2021). Development of mathematics learning tools with realistic mathematics education-jumping task (RME-JT) and its effect on the mathematic communication skills. *Journal of Physics: Conference Series, 1839*(1), Article 012018. https://doi.org/10.1088/1742-6596/1839/1/012018

Sagita, L., Putri, R. I. I., & Prahmana, R. C. I. (2022). Promising research studies between mathematics literacy and financial literacy through project-based learning. *Journal on Mathematics Education, 13*(4), 753–772. https://doi.org/10.22342/jme.v13i4.pp753-772

Saleh, M., Prahmana, R. C. I., Isa, M., & Murni, M. (2018). Improving the reasoning ability of elementary school students through the Indonesian realistic mathematics education. *Journal on Mathematics Education, 9*(1), 41–54. https://doi.org/10.22342/jme.9.1.5049.41-54

Suherman, S., & Vidákovich, T. (2022a). Adaptation and validation of students' attitudes toward mathematics to Indonesia. *Pedagogika, 147*(3), 227–252. https://doi.org/10.15823/p.2022.147.11

Suherman, S., & Vidákovich, T. (2022b). Assessment of mathematical creative thinking: A systematic review. *Thinking Skills and Creativity, 44*, Article 101019. https://doi.org/10.1016/j.tsc.2022.101019

Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st-century skills for the future of work and education. *Journal of Intelligence, 11*(3), Article 54. https://doi.org/10.3390/jintelligence11030054