Indonesian Technology and Education Journal
E-ISSN: 2985-8216; P-ISSN: 2985-8208
https://journal.diginus.id/index.php/ITEJ

Google Sites-Mediated Problem-Based Learning: Effects on Students' Conceptual Understanding and Mathematical Self-Efficacy

Risna Helmi[1], Yumiati[2]*, Endang Wahyuningrum[3]
[1,2,3] *Master's Program in Mathematics Education, Graduate School, Universitas Terbuka, Tangerang Selatan, Indonesia*
**Correspondence: yumi@ecampus.ut.ac.id*

ARTICLE INFO

Keywords:

Conceptual Understanding; Geometric Transformation; Google Sites; Mathematical Self-Efficacy; Problem-Based Learning; 4D Model.

Article History

Received: Aug 15, 2026
Revised: Sept 20. 2026
Accepted: Oct 22, 2026

ABSTRACT

Background: Geometric transformations require students to connect visual, coordinate, symbolic, and matrix representations. In the participating class, procedural answers frequently concealed incomplete conceptual explanations, while many students hesitated to tackle unfamiliar tasks. An integrated digital environment was therefore designed to support the phases of problem-based learning (PBL) and opportunities to develop mathematical self-efficacy.

Aim: This study developed a PBL-based Google Sites environment for geometric transformations and examined its validity, practicality, and pre–post changes in conceptual understanding and self-efficacy.

Method: Research and development followed the Define, Design, Develop, and Disseminate (4D) model. A content expert, media expert, and practitioner validated the product. Its field trial involved 21 Grade XI students in Batam, Indonesia, over six PBL meetings using a one-group pretest–posttest design. Evidence came from validation and response forms, implementation observation, a ten-item conceptual understanding essay test ($\alpha = .874$ reported for the instrument try-out), and a 16-item self-efficacy scale ($\alpha = .93$ before and .85 after). Percentage scores, normalized gain, and Wilcoxon signed-rank tests were used.

Results: The environment obtained 88.61% for validity (content 93.08%, media 80.77%, practitioner 93.00%) and 93.97% for practicality (implementation 92.22%, teacher response 100%, student response 89.68%). Conceptual understanding increased from M = 66.52 to M = 87.81 (mean N-gain = 0.67; mastery 28.57% to 100%; $z = 4.02$, $p < .001$, $r = .88$). Self-efficacy increased from M = 45.43 to M = 59.24 (mean N-gain = 0.81; $z = 4.02$, $p < .001$, $r = .88$), with the largest increase in magnitude.

Conclusion: The PBL-aligned website was rated valid and practical, and both measured outcomes improved during the field trial. The absence of a comparison group limits causal attribution and wider generalization.

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

To cite this article : Helmi, R., Yumiati, Y., & Wahyuningrum, E. (2026). Developing a PBL-Based Google Sites Learning Environment for Geometric Transformations: Validity, Practicality, and Pre–Post Changes in Conceptual Understanding and Mathematical Self-Efficacy. *Information Technology Education Journal* 4(3), 12-33. Doi. 10.61255/itej.v4i3.2400

INTRODUCTION

Conceptual understanding in mathematics concerns the meaning, structure, and relationships of ideas, rather than the mechanical reproduction of procedures (Kilpatrick et al., 2001; Yang et al., 2021). It enables learners to coordinate prior knowledge, select representations, and explain why a procedure works in a new situation (Jäder & Johansson, 2025). Recent work on conceptual and procedural knowledge shows that their relationship is dynamic, while an apparently correct answer may still conceal an unresolved misconception (Lenz et al., 2024; Uegatani et al., 2024). Classroom feedback and the connections teachers make between

representations therefore matter for how students interpret mathematical objects (Dreher et al., 2024; Stovner & Klette, 2022). Digital tools can expand representational access, but that access requires pedagogical guidance rather than assuming that visualization itself establishes understanding (Bach et al., 2024; Engelbrecht & Borba, 2024).

Geometric transformation makes these challenges visible because translation, reflection, rotation, dilation, and composition involve changes that must be interpreted across diagrams, coordinates, rules, and matrices. Students may correctly compute an image point without articulating the invariant properties of the transformation or the meaning of the displacement vector. In the participating Grade XI class, only 6 of 21 students (28.57%) produced complete explanations on a preliminary translation task, whereas 15 (71.43%) omitted reasoning, reversed the direction of a shift, or could not transfer the vector to another point. These difficulties are relevant to research in which coordinated visual activity and mathematical discussion support geometry learning, although the effectiveness of particular tools depends on instructional design (Gurmu et al., 2024; Valori et al., 2024). The local diagnosis consequently called for activities that made students' explanations and representational choices visible before procedural rules were consolidated.

Conceptual difficulties were accompanied by a related affective concern: 14 of the 21 students (66.67%) hesitated during problem solving or avoided explaining an unfamiliar solution. Mathematical self-efficacy refers to learners' beliefs about their capability to carry out mathematical tasks, including when those tasks become demanding (Bandura, 1997; Street et al., 2024). It is associated with persistence and achievement, while its relationship to mathematics anxiety and performance is neither trivial nor uniform across learners (Holenstein et al., 2022; Živković et al., 2023; Sakellariou, 2022; Shimizu, 2025). Studies of the sources, measurement, and development of such beliefs point to the importance of task specificity, opportunities for successful performance, feedback, and emotional conditions (Herset et al., 2023; Koponen et al., 2021; Yu et al., 2023; Zakariya, 2021, 2022). The present design therefore treated confidence as an instructional concern alongside conceptual understanding, rather than an incidental benefit of digital media.

PBL offers a sequence in which an authentic problem precedes investigation, the comparison of solutions, and reflection (Arends, 2012; Hmelo-Silver, 2004). Its collaborative and exploratory activities create opportunities to explain mathematical decisions while experiencing success, peer modeling, and constructive feedback. Reviews and classroom investigations have associated problem-oriented instruction with mathematical reasoning and higher-order thinking, although outcomes depend on the structure of tasks and the duration and quality of implementation (Aba-Oli et al., 2025; Hartmann et al., 2021; Negara et al., 2023; Nst et al., 2023; Sudianto & Ramdiani, 2024). Digital supports may reduce the organizational burden of coordinating problems, visual examples, worksheets, and reflection, but technology should be integrated into the pedagogy rather than added as an isolated repository (Cevikbas & Kaiser, 2023; Hidayat & Firmanti, 2024). Google Sites was selected because it could bring the school's existing worksheets, videos, formative tasks, and instructions into one accessible sequence (Waseso et al., 2022).

Existing Google Sites development studies report favorable validation and usability results across mathematics topics (Ciung et al., 2022; Kristanto & Afifah, 2022). Some have examined conceptual understanding, problem solving, or affective dispositions (Jannah et al., 2024; Johdi et al., 2024; Noverlika et al., 2024; Rahmafani & Sulistyaningrum, 2025; Riyanto et al., 2025). More specifically, Mustika et al. (2024) examined PBL supported by Google Sites in quadratic functions, and other classroom reports address comparable mathematical outcomes (Pakih et al., 2025). These studies justify investigating an integrated website, but favorable media ratings are not equivalent to evidence that a particular website feature supports a particular PBL process.

International research on digital mathematics learning likewise stresses the need to distinguish the availability of multiple representations from the explanations and interactions that make them educationally useful (Bach et al., 2024; Dreher et al., 2024; Engelbrecht & Borba, 2024).

Three connected gaps frame the study. First, the relationship between a site's individual components and the five PBL phases is often less explicit than the description of its technical features. Second, research on mathematical conceptual understanding and on self-efficacy commonly addresses the outcomes separately, even though a review suggests theoretically relevant links between them (Fitrianingsih et al., 2024; Street et al., 2024). Third, Google Sites development reports have frequently focused on other school topics, whereas senior-secondary geometric transformations require coordinated visual and symbolic reasoning (Arnisya et al., 2025; Kristanto & Afifah, 2022). Research on guided mathematical tasks also cautions that choice, task complexity, and peer participation do not improve motivational beliefs automatically (Herset et al., 2023; Krawitz et al., 2026). Together, these considerations motivated a design that specified instructional functions for site components and examined cognitive and affective outcomes in the same field trial.

This study developed and tested a PBL-based Google Sites environment for geometric transformations using the four-stage 4D model. Its central contribution is a design matrix that connects each PBL phase to a website component, a possible source of self-efficacy, and an indicator of conceptual understanding; this is a pedagogical design proposition rather than an experimentally established mechanism. Four questions guided the investigation: how the environment was developed; how content, media, and practitioner experts assessed its validity; how classroom observation and user responses characterized practicality; and how students' conceptual understanding and self-efficacy changed from pretest to posttest. The evaluation used one intact class, so it assessed feasibility and promising change under local implementation conditions rather than a comparison with alternative instruction. This bounded contribution may help mathematics teachers adapt a coherent sequence of problems, resources, feedback, and reflection to the representational demands of transformation geometry.

THEORETICAL FRAMEWORK

Conceptual understanding of geometric transformations

Following Kilpatrick et al. (2001) and indicators used in mathematics learning assessments (Afifah et al., 2024; Hardita et al., 2024), conceptual understanding was represented by six processes: restating a concept; classifying objects by properties; giving examples and nonexamples; moving between representations; selecting and using appropriate procedures; and applying concepts in problem solving. In geometric transformations, these processes include explaining why $A(x, y)$ becomes $A'(x + a, y + b)$ under $T(a, b)$, identifying the meaning of a reflection line, interpreting rotation in graphical and matrix form, and reasoning about the order of composite transformations. Research on mathematical representations suggests that the explanatory links between drawings and symbols, rather than the mere availability of diagrams, deserve explicit instructional attention (Bach et al., 2024; Dreher et al., 2024; Gurmu et al., 2024). The test operationalized these six processes through ten assessment targets distributed across the lesson content (Table 4).

Mathematical self-efficacy

The self-efficacy scale used Bandura's (1997) dimensions of magnitude, strength, and generality, referring respectively to confidence across task difficulty, stability of the belief when obstacles arise, and perceived applicability across mathematical situations. Four proposed sources of efficacy beliefs are mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states (Usher & Pajares, 2008). Recent studies underscore that how

tasks are sequenced, the opportunity to evaluate progress, and students' interpretations of successful or difficult experiences can matter for confidence (Herset et al., 2023; Koponen et al., 2021; Vogelsanger-Holenstein et al., 2025; Zakariya, 2022). These sources informed the design of LKPD tasks, group presentation, feedback, and reflection; their individual contributions were not measured in the present study.

Design conjecture: PBL syntax mediated by Google Sites

The design conjecture was that Google Sites could make the five PBL phases visible and navigable: orientation to a contextual problem, organization for learning, guided investigation, development and presentation of work, and evaluation of problem-solving processes (Arends, 2012). Each phase was linked to a particular website component and to intended opportunities for conceptual explanation and self-efficacy. The design drew on evidence that digital and collaborative environments can support mathematical discussion when tasks and interactions are deliberately structured (Giberti et al., 2022; Tempelaar et al., 2024). Figure 2 displays this proposed alignment; the field trial evaluates the integrated environment and does not isolate the effect of any single website component.

METHOD

Research design

The study adopted research and development guided by the 4D model: Define, Design, Develop, and Disseminate (Thiagarajan et al., 1974). The model organized the identification of instructional needs, the specification of a PBL-aligned prototype, expert review and revision, a field trial, and limited dissemination within the participating school (Sihombing et al., 2024). Within the Develop stage, pretest and posttest measures were administered to the same students, without random assignment or a separate comparison class. This design is suitable for describing change during initial product implementation, but it cannot distinguish the intervention from maturation, repeated testing, or other concurrent influences. Figure 1 identifies the sequence and corresponding output of each development stage. The complete development sequence and its stage-specific outputs are displayed in Figure 1.

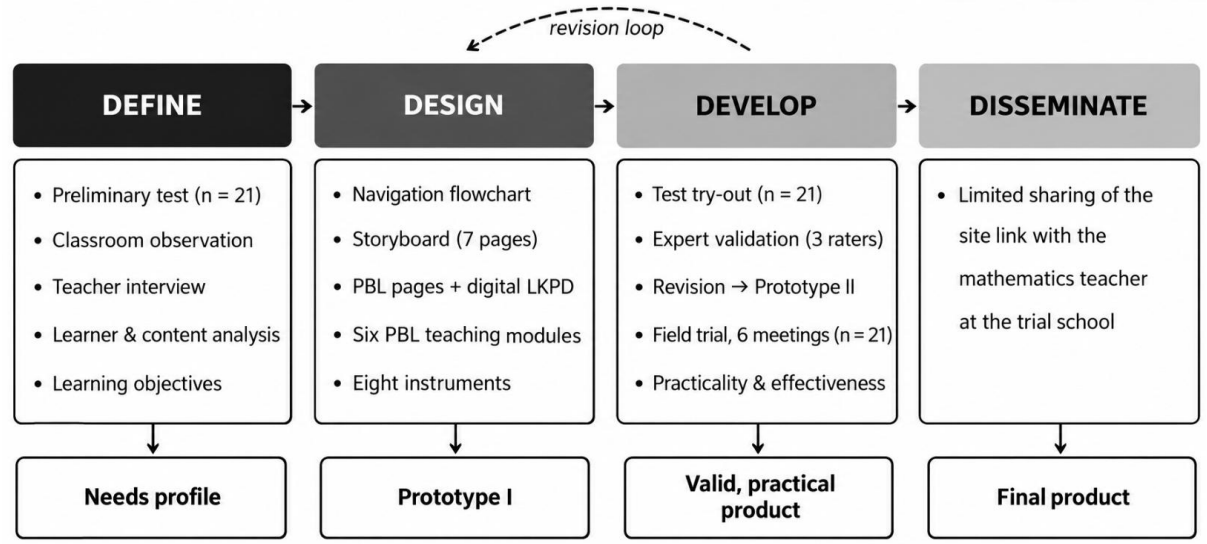


Figure 1. Research procedure based on the 4D model and the output of each stage

The flowchart shows how the preliminary diagnosis informed a prototype that was then validated, revised, and trialed before limited dissemination. It represents the procedure followed rather than a controlled comparison.

Setting and participants

The investigation took place at a public senior high school in Batam, Riau Islands Province, Indonesia, from April to June 2026, under the Merdeka Curriculum (Phase F). The trial involved one intact Grade XI advanced mathematics class (XI-2) comprising 21 students, purposively selected because the class was scheduled to study geometric transformations and had adequate internet access. The class teacher served as observer and completed the teacher-response questionnaire. An independent group of 21 Grade XI students outside the trial class participated in the test-item try-out. Three people reviewed the website: a lecturer in mathematics education as content expert, a lecturer specializing in digital learning media as media expert, and an experienced mathematics teacher as practitioner. University and school permissions were obtained, and participants were represented in the dataset by anonymized codes S-01 to S-21.

Instruments

Eight instruments were prepared during the Design stage. The conceptual understanding test initially included 20 essay items representing ten indicator-by-content assessment targets, with two items per target; these targets operationalized the six conceptual processes defined in the theoretical framework. In the independent try-out (n = 21), 14 of the 20 items met the reported Pearson product–moment threshold (r-table = .433, df = 19, α = .05); 19 items were of moderate difficulty and one was difficult, while item discrimination ranged from very good (8 items) to very poor (4 items). Ten valid items with acceptable discrimination were retained for pretest and posttest, yielding a maximum raw score of 40 before conversion to a 0–100 scale. The instrument try-out reported Cronbach's α = .874. The 16-item self-efficacy scale comprised magnitude (5 items), strength (5 items), and generality (6 items); four negatively worded items were reverse-scored, giving a total range of 16–64. Its internal-consistency coefficients were α = .93 at pretest and α = .85 at posttest. The remaining instruments were content (26 items), media (26 items), and practitioner (20 items) validation sheets using five-point scales; a four-point implementation observation sheet comprising 25 indicators across lesson opening, PBL phases, and closure; and teacher (15 items) and student (18 items) response questionnaires using five-point scales. Non-test instruments were subjected to expert content review before administration. Evidence on measurement dimensionality and task-specific efficacy assessment motivates cautious interpretation of totals and dimensions (Navarro et al., 2025; Yu et al., 2023).

Data analysis

Product validity and practicality were expressed as a percentage of the respective maximum possible score, P = (obtained score / maximum score) × 100%. The descriptive interpretation used the study's predetermined bands: 81.00–100% very valid/practical, 61.00–80.99% valid/practical, 41.00–60.99% fairly valid, 21.00–40.99% less valid, and ≤20.99% invalid; the minimum acceptance threshold was the valid/practical category. Individual learning change was summarized using normalized gain, g = (posttest − pretest) / (maximum − pretest), with high ($g \geq .70$), medium ($.30 \leq g < .70$), and low ($g < .30$) classifications (Hake, 1998). Conceptual mastery was assessed using the school's learning-objective criterion (KKTP) of 75. Because paired differences in conceptual understanding did not meet the Shapiro–Wilk normality criterion (W = .872, p = .010), two-tailed Wilcoxon signed-rank tests were used for paired outcomes, including self-efficacy totals and dimensions. Effect size was expressed as $r = Z/\sqrt{N}$, with $r \geq .50$ classified as large (Fritz et al., 2012), and statistical significance was judged at α = .05. Under the study's operational criterion, a medium-or-higher mean N-gain together with a significant paired

difference was treated as evidence of promising effectiveness during the trial, not proof of a causal intervention effect.

RESULTS AND DISCUSSION

The results are presented in the order of the 4D stages. Each stage is reported together with the evidence that informed the next stage, so that the development decisions can be traced from the initial needs to the final product.

A. Define stage: establishing the instructional need

The Define stage comprised five analyses: a front-end diagnosis of students' conceptual understanding and self-efficacy, classroom observation, a semi-structured teacher interview, learner analysis, and concept and task analysis. Each analysis is reported below, and their combined implications are synthesised at the end of this subsection as explicit design requirements.

1. Front-end diagnosis

Table 1 presents the diagnostic evidence that motivated the study. On two essay items on translation, fewer than one third of the students produced complete and correctly reasoned answers. The dominant error patterns, writing only final answers, reversing the direction of the coordinate shift, and failing to apply the same vector to a new point, indicate a conceptual rather than merely computational difficulty. A parallel pattern appeared in the affective domain: two thirds of the students displayed behaviors associated with low self-efficacy during problem solving.

Table 1. Front-end diagnosis of conceptual understanding and self-efficacy (n = 21)

Domain	What was assessed	Result n (%)	Observed pattern
Conceptual understanding (two translation items)	Determining a translation vector, explaining its meaning, using coordinate representation, applying it to another point	Complete: 6 (28.57) Incomplete: 15 (71.43)	Final answers without reasoning; reversed direction of shift; inability to explain the vector; failure to transfer the vector to a new point
Self-efficacy (classroom observation)	Confidence when solving and explaining problems in class	Confident: 7 (33.33) Hesitant: 14 (66.67)	Hesitation when starting tasks; reluctance to explain answers in front of the class; giving up on tasks that differ from worked examples

The diagnosis indicates that incomplete mathematical reasoning and visible hesitation co-occurred in this class. These observations established the two learning outcomes targeted by the prototype.

2. Classroom observation and teacher interview

Classroom observation (Table 2) showed that lessons were well organized and that digital facilities were available, yet the media used by the teacher were scattered across separate platforms. The teacher interview (Table 3) confirmed this picture and identified the subtopics that students found most difficult.

Table 2. Results of the preliminary classroom observation

Aspect observed	Finding
Course of the lesson	Lesson followed the planned sequence: objectives, explanation, worked examples, group discussion, exercises, and conclusion
Learning media	Textbook, printed worksheets (LKPD), PowerPoint slides, and instructional videos
Teacher activity	Actively guided students, invited questions, managed discussion, and gave feedback
Student activity	Most students took part in discussion, but some lost focus and lacked confidence when presenting
Digital facilities	LCD projector and internet available; most students owned a smartphone
Overall conclusion	Learning was well run and digitally supported, but media were fragmented and not integrated in one platform

The classroom already had basic digital facilities, but lesson resources were distributed across different media. Integration was therefore selected as a practical design requirement. The teacher interview provides an additional perspective on content difficulty and media needs, as outlined in Table 3.

Table 3. Summary of the teacher interview and its implications

Aspect	Teacher's account	Implication for design
Instructional model	Direct instruction, group discussion, and PBL for selected topics	PBL is familiar to the teacher and can be implemented consistently
Geometric transformation	Rotation, dilation, and composite transformations are the most difficult subtopics	Richer visual support and scaffolded prompts for these subtopics
Conceptual understanding	Many students memorise formulae without understanding the underlying concept	Start each topic with a meaning-oriented problem before rules
Self-efficacy	Many students lack confidence, hesitate to express opinions, and give up easily	Graded tasks, group work, and supportive feedback
Media in use	Textbook, LKPD, PowerPoint, Google Classroom, Google Forms, YouTube	Integrate existing tools into one site rather than adding new ones
Need for Google Sites	Supports development because it can integrate materials, LKPD, and evaluation	Teacher acceptance supports practicality

The teacher identified rotation, dilation, and composition as demanding subtopics and supported a site-based PBL sequence. These perspectives informed the choice of content and scaffolds.

3. Learner analysis and concept–task analysis

The learner analysis showed that students' academic ability was heterogeneous, that most were accustomed to using smartphones and other digital devices, and that their active engagement and confidence in mathematics still needed to be strengthened. A mobile-accessible web environment was therefore judged appropriate. The concept and task analysis, based on the Phase F learning outcomes of the Merdeka Curriculum, organized the content into six meetings. Table 4 shows how each meeting was specified in terms of learning objectives, the conceptual understanding indicators later assessed by the test, and the difficulties anticipated from the diagnosis and the teacher interview.

Table 4. Concept and task analysis: specification of meetings, objectives, and assessed indicators

Meeting / subtopic	Main learning objectives	Assessed indicator of conceptual understanding	Anticipated difficulty
M1 Basic concept	Explain transformation in one's own words; identify object and image; recognise changes in position, orientation, shape, and size	Restating the concept of geometric transformation	Distinguishing the four types of transformation
M2 Translation	Determine images of points and shapes; determine the translation vector from an object and its image	Explaining translation through change of position; using translation to find images	Direction of the shift; transferring the vector to other points
M3 Reflection	Discover and use rules for reflection in the x-axis, y-axis, origin, y = x, and y = −x	Classifying reflections by mirror line; giving examples and non-examples	Signs of the image coordinates
M4 Rotation	Identify centre, angle, and direction; determine images, including by matrix	Representing rotation in coordinate or graphical form; using the rotation matrix	Direction and size of the angle
M5 Dilation	Identify centre and scale factor; determine images; analyse the effect of k	Using dilation to determine images	Centre of dilation not at the origin
M6 Composite	Determine images under successive transformations using matrices; solve contextual problems	Selecting procedures for composite transformations; applying transformations in context	Order of the transformations

The six-meeting specification connects subtopic objectives to targeted conceptual processes and anticipated difficulties, providing a basis for designing the LKPD and assessments.

4. Synthesis: from needs to design requirements

Table 5 condenses the Define stage into five design requirements. Each requirement links an identified need to a concrete feature of the product, which provides an auditable bridge between needs analysis and design.

Table 5. Design requirements derived from the Define stage

No	Identified need	Design requirement	Realised in
R1	Rote use of formulae; weak explanation of meaning (Tables 1, 3)	Begin every meeting with a contextual problem that asks what changes and what stays invariant	Learning page and orientation tasks in the LKPD
R2	Fragmented media across platforms (Tables 2, 3)	Integrate materials, LKPD, evaluation, and reflection in one environment	Single Google Sites environment with fixed navigation
R3	Difficulty with rotation, dilation, and composition (Tables 3, 4)	Provide multiple representations and scaffolded prompts	Images, videos, worked examples, and guiding questions
R4	Low confidence; giving up on unfamiliar tasks (Table 1)	Sequence tasks from simple to complex and make peer success visible	Graded LKPD tasks, group roles, and presentation phase
R5	Heterogeneous ability; smartphone familiarity (learner analysis)	Allow self-paced revisiting outside class on mobile devices	Openly accessible, mobile-friendly pages; Guide page

The five requirements trace design choices directly to observed needs. This traceability provides a rationale for the product features without claiming that each feature independently produced learning gains.

B. Design stage: constructing the PBL-aligned prototype

The Design stage translated the five requirements into a design matrix, a navigation structure and storyboard, the website pages, six teaching modules, and eight research instruments. These outputs are reported in turn.

1. Design matrix

The central design decision was to assign every page of the site a role in a specific PBL phase and a deliberate function in building self-efficacy and conceptual understanding, rather than adding PBL activities to a generic website (Figure 2). For example, the Learning page opens each meeting with a contextual problem, which lowers the threshold of entry (affective states) and asks students to restate and represent the concept, while the embedded digital LKPD sequences tasks from simple to complex so that students accumulate mastery experiences during guided investigation. Figure 2 illustrates how the product design connected PBL activities with specific pages and intended learning targets.

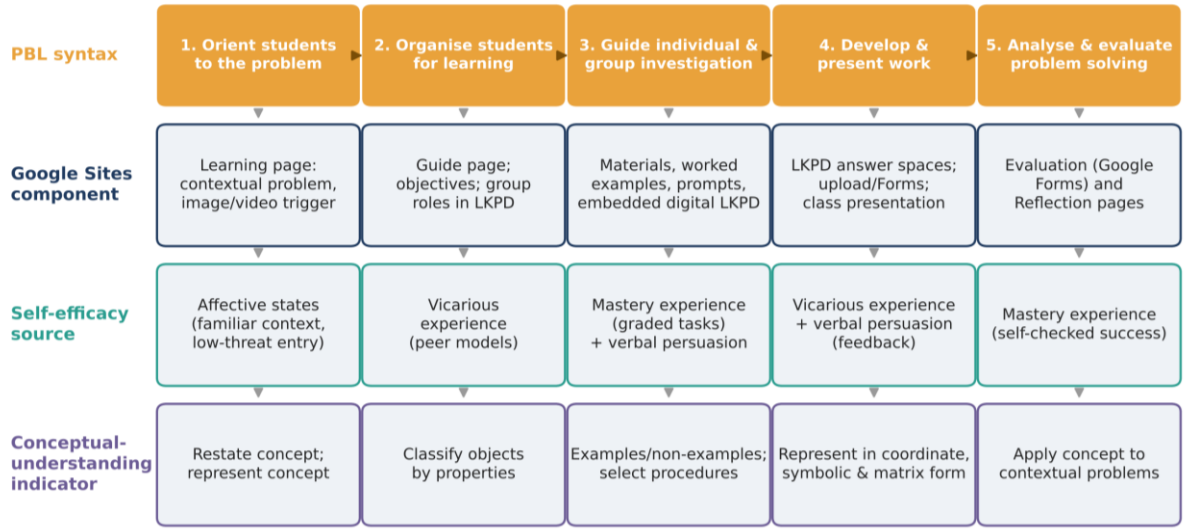


Figure 2. Design matrix aligning the PBL syntax with Google Sites components, sources of self-efficacy, and indicators of conceptual understanding

The matrix makes the instructional purpose of each site component explicit. Its proposed self-efficacy sources are theoretical design targets rather than separately tested mediators.

2. Navigation structure, storyboard, and pages

The navigation flow was linear but open: students move from the Home page to the Guide, the Learning menu, the LKPD, Evaluation, Reflection, and the developer Profile, while still being able to revisit any page outside class. Table 6 presents the storyboard together with the design requirement that each page addresses, and Figure 3 shows the main pages of the prototype.

Table 6. Storyboard of the Google Sites environment

Page	Content	Purpose	Requirement
Home	Title, illustrations, brief description, button to the guide	Orient students to the environment and the learning flow	R2
Guide	Step-by-step instructions for using the site	Reduce navigational uncertainty and anxiety	R4, R5
Learning	Materials on the basic concept, translation, reflection, rotation, dilation, and composition, with contextual problems, images, videos, and worked examples	Build concepts progressively from contextual problems	R1, R3

Page	Content	Purpose	Requirement
LKPD	Digital PBL worksheet embedded for each meeting	Guide investigation, group work, and presentation of solutions	R1, R4
Evaluation	Google Forms practice for each meeting	Give feedback on conceptual understanding	R2, R4
Reflection	Prompts on understanding, difficulties, helpful strategies, and confidence	Support self-assessment of process and beliefs	R4
Profile	Developer identity	Inform users about the developer	–

The storyboard assigns a specific instructional role to each page. Students could navigate the sequence while returning to earlier materials when additional explanation was needed. The main interface and embedded learning activities are illustrated in Figure 3.

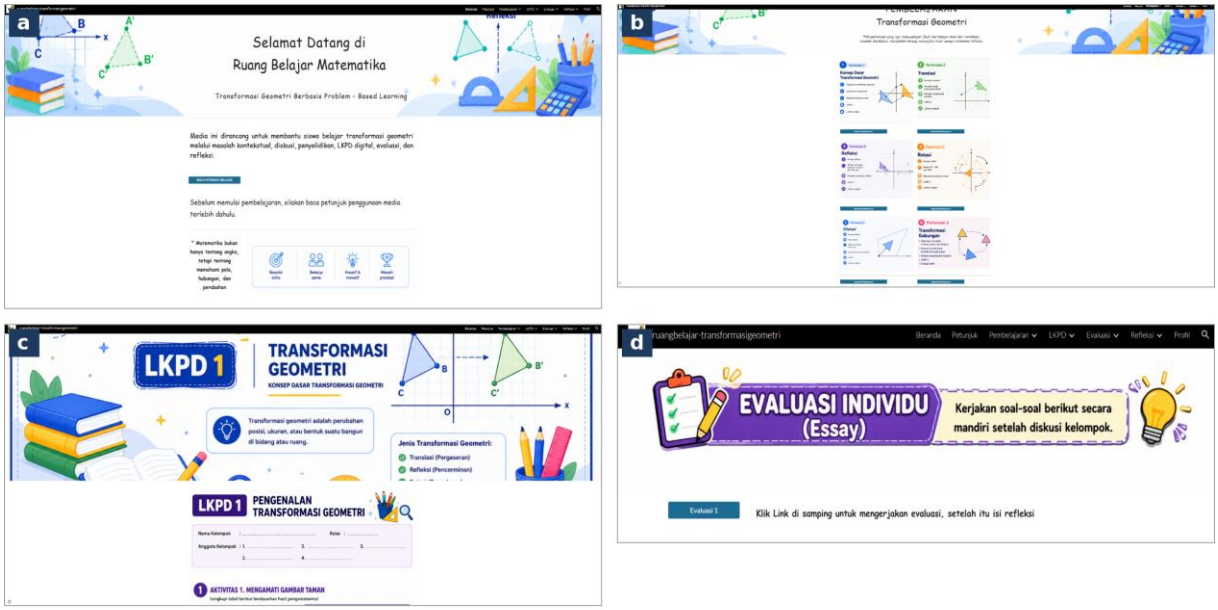


Figure 3. Main pages of the prototype: (a) Home, (b) Learning materials, (c) digital PBL worksheet (LKPD), and (d) Evaluation

The examples show the Home page, the Learning menu, a digital LKPD, and an evaluation activity. Together, they illustrate integration into one navigation environment.

3. Teaching modules

Six teaching modules of 3 × 45 minutes were prepared. Each module followed the same internal structure (introduction about 15 minutes, five PBL phases about 105 minutes, closure about 15 minutes) and linked directly to the corresponding pages, so that the LKPD was an integral part of the site rather than a separate printed document. Table 7 summarizes how the PBL phases were enacted through the site in each meeting.

Table 7. Structure of the PBL teaching modules enacted through Google Sites

Meeting	Orientation problem on the site	Investigation in the digital LKPD	Concept consolidated in phase 5
M1	Observing a garden scene in which objects change position, orientation, or size	Identifying objects and images; describing what changes	Definition of transformation; first look at the four types

Meeting	Orientation problem on the site	Investigation in the digital LKPD	Concept consolidated in phase 5
M2	A contextual situation in which an object is moved to a new position	Tracing coordinate changes; finding the vector from object and image	T(a, b): (x, y) → (x + a, y + b)
M3	An object and its mirror image	Investigating coordinate patterns for five mirror lines	Rules for reflection in the axes, origin, y = x, and y = −x
M4	An object turning about a point	Investigating centre, angle, and direction; using the rotation matrix	Relationship between centre, angle, direction, and image
M5	An object being enlarged or reduced	Exploring the effect of the scale factor on size and position	Dilation rule with centre O and factor k
M6	A situation requiring successive transformations	Applying transformations in order using matrices	Composite transformations and the role of order

Across all meetings, contextual orientation problems preceded investigation and evaluation. The progressive sequence covered basic transformation ideas through compositions.

4. Research instruments

The Design stage also produced eight instruments (Table 8). The conceptual understanding test was the only instrument that required empirical try-out; the other seven were content-validated by experts before use.

Table 8. Research instruments developed in the Design stage

Instrument	Aspects / dimensions	Items	Scale	Respondent
Content validation sheet	Content, presentation, language, PBL syntax, CU indicators	26	1–5	Content expert
Media validation sheet	Consistency, language, graphics, ease of use	26	1–5	Media expert
Practitioner validation sheet	Fit with learning, PBL, ease of use, usefulness, feasibility	20	1–5	Teacher
Implementation observation sheet	Opening, five PBL phases, closing	25	1–4	Observer
Teacher response questionnaire	Content, display, PBL alignment, usefulness	15	1–5	Teacher
Student response questionnaire	Content, display, ease of use, usefulness	18	1–5	Students
Conceptual understanding test	Ten assessment targets from six indicators (Table 4)	10 essay	0–4 per item	Students
Self-efficacy scale	Magnitude, strength, generality	16	1–4	Students

The instrument inventory separates validation, implementation, user response, conceptual understanding, and self-efficacy measures. Test targets and theoretical indicators refer to different levels of assessment design.

C. Develop stage: validation, revision, and field trial

1. Validity

Table 9 summarizes the validation results. The overall score of 88.61% places the environment in the very valid category. The content expert rated the product very valid (93.08%), with perfect alignment with the conceptual understanding indicators (100%) and high scores for

content feasibility (95.00%), language (95.00%), alignment with the PBL syntax (92.00%), and presentation (86.67%). The practitioner also rated it very valid (93.00%). The media expert gave the lowest score (80.77%, valid), and the aspect profile in Figure 4 shows why: consistency (mean 4.33 of 5) and language (4.25) were strong, whereas graphics (3.60) and ease of use (3.67) were the weakest aspects across all validators.

Table 9. Expert and practitioner validation of the Google Sites environment

Validator	Score obtained	Maximum score	Percentage	Category
Content expert	121	130	93.08%	Very valid
Media expert	105	130	80.77%	Valid
Practitioner	93	100	93.00%	Very valid
Overall	**319**	**360**	**88.61%**	**Very valid**

The aggregate validation percentage was 88.61%. The media score was the lowest of the three assessments, which motivated revisions to graphics and guidance. Figure 4 compares the aspect-level ratings supplied during expert and practitioner validation.

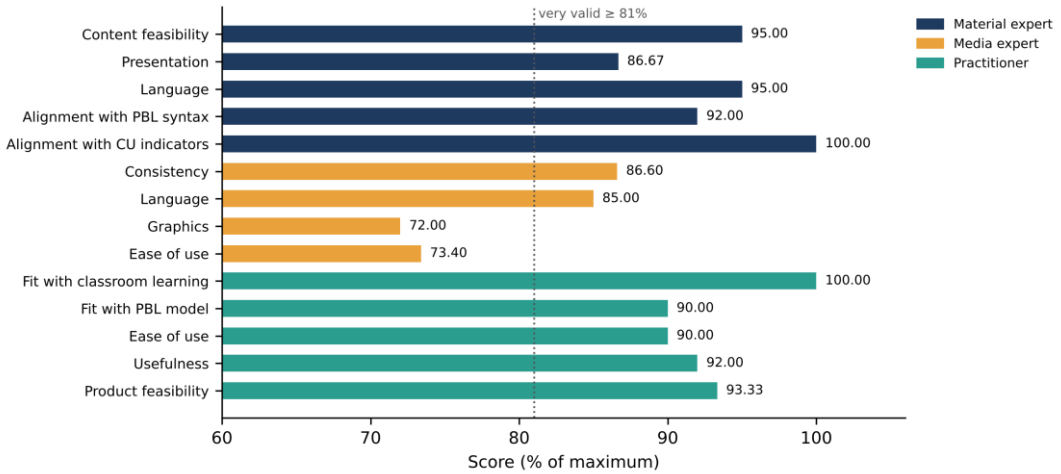


Figure 4. Validation profile by aspect (media expert means on the 1–5 scale converted to percentages)

The profile shows stronger content alignment than graphical presentation and ease of use. The latter two aspects shaped revisions before the classroom trial. The validators' comments were consistent with the aspect profile and were translated into three concrete revisions, summarized in Table 10. The revised version (Prototype II) was judged suitable for the field trial.

Table 10. Revisions made on the basis of validators' suggestions

Revised component	Before revision	After revision
Materials and LKPD	Several sections dominated by images without accompanying text	Explanatory captions added to every image that carries a concept
Usage guide	Presented mainly as screenshots	Short, numbered textual steps added alongside the screenshots
Evaluation page	Typographical errors in the instructions	Errors corrected and instructions re-checked

The revisions addressed readability, navigational instructions, and errors in the evaluation page. Their completion was documented, although expert re-rating after revision was not reported.

2. Practicality

The field trial took place over six meetings, preceded by the pretest and the initial self-efficacy questionnaire and followed by the posttest, the final questionnaire, and the response questionnaires. Practicality was evaluated from three sources (Table 11). The observer recorded that 83 of the 90 planned activities were carried out (92.22%, very good), with phases 1–3 fully implemented and phases 4 and 5 at 91.67% each. The teacher gave the maximum score on all 15 statements (100%), and the 21 students gave a total of 1,695 of 1,890 points (89.68%), with individual scores ranging narrowly from 88.89% to 90.00%. The mean practicality was 93.97%, which falls in the very practical category.

Table 11. Practicality of the Google Sites environment

Data source	Respondents	Percentage	Category
Implementation observation (6 meetings)	1 observer	92.22%	Very good
Teacher response questionnaire	1 teacher	100.00%	Very practical
Student response questionnaire	21 students	89.68%	Very practical
Mean		**93.97%**	**Very practical**

The meeting-by-meeting ratings on the four-point observation scale provide a more dynamic picture (Figure 5). The implementation index rose from 87% in the first meeting to 92% in the second and stabilized at 94% from the third meeting onwards. The largest improvements occurred in the phases that depend most on students' autonomy: analyzing and evaluating the problem-solving process increased from 83% to 100%, and developing and presenting work increased from 75% to 85%. The observer's notes confirm that students initially needed guidance in navigating the site and in presenting their solutions but became increasingly independent as they grew familiar with the environment. The presentation phase nevertheless remained the least fully implemented, mainly because of time constraints. The meeting-by-meeting implementation pattern is presented in Figure 5.

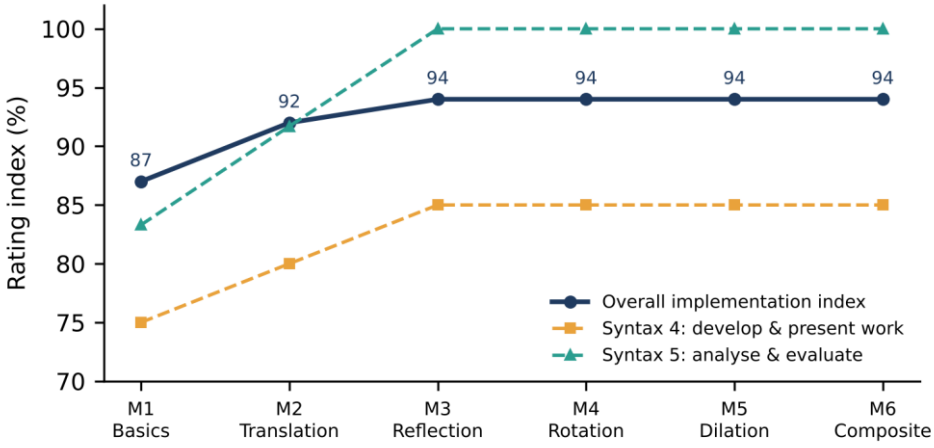


Figure 5. Implementation of the PBL syntax across the six meetings (observer ratings expressed as a percentage of the maximum)

Recorded ratings increased over the first meetings and then stabilized; presentation and evaluation activities remained less consistently enacted than earlier PBL phases.

3. Effectiveness for conceptual understanding

Table 12 reports the descriptive and Wilcoxon statistics for conceptual understanding, while Figure 6a displays individual score trajectories. The mean score increased from 66.52 (SD

= 14.17) to 87.81 (SD = 7.72), an observed difference of 21.29 points; the minimum increased from 43 to 80. All 21 paired differences were positive (W = 0, z = 4.02, p < .001, r = .88). The mean individual N-gain was 0.67 (medium): six students (28.57%) were in the high category and 15 (71.43%) in the medium category. The proportion reaching the KKTP threshold of 75 increased from 6 students (28.57%) to 21 (100%). The smaller posttest standard deviation indicates reduced dispersion, but does not by itself establish which baseline subgroup benefited most.

Table 12. Descriptive and inferential statistics for conceptual understanding (n = 21)

Statistic	Pretest	Posttest	Difference / test
Mean (SD)	66.52 (14.17)	87.81 (7.72)	+21.29
Median	68.00	85.00	+17.00
Minimum – maximum	43 – 88	80 – 100	
Students reaching KKTP ≥ 75	6 (28.57%)	21 (100%)	+15 students
Mean N-gain (category)			0.67 (medium)
Wilcoxon signed-rank			W = 0; z = 4.02; p < .001
Effect size			r = .88 (large)

Individual pretest and posttest outcomes for both measures are visualized in Figure 6.

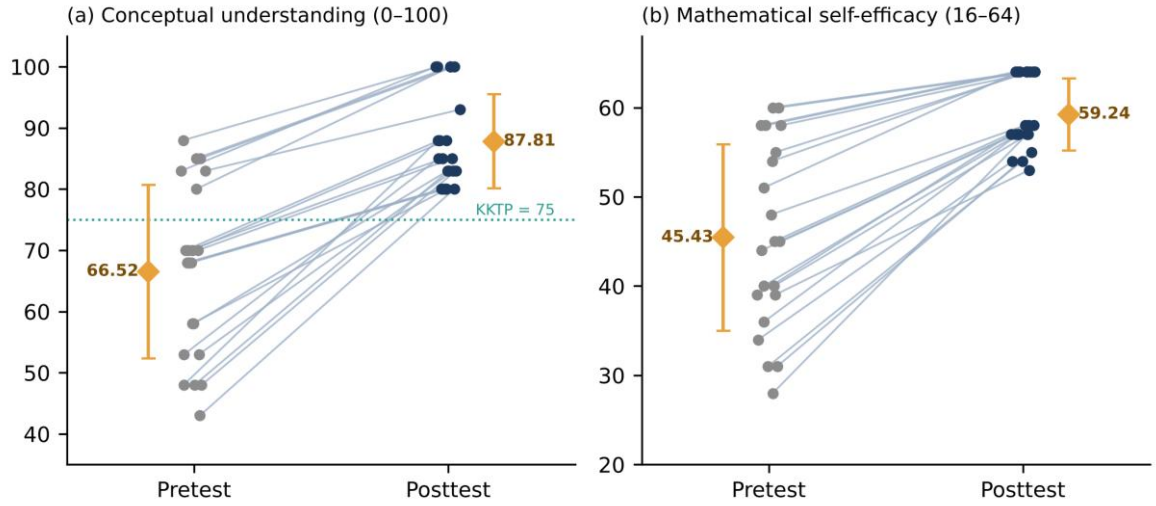


Figure 6. Individual pretest–posttest trajectories for (a) conceptual understanding and (b) mathematical self-efficacy; diamonds show means ± 1 SD

The trajectories illustrate the generally upward paired pattern and the narrower posttest spread. They do not identify which part of the intervention accounts for the change.

4. Effectiveness for mathematical self-efficacy

Self-efficacy scores also increased between the two measurement points (Table 13; Figure 6b). The total mean rose from 45.43 (SD = 10.45) to 59.24 (SD = 4.05), and the minimum increased from 28 to 53. Each of the 21 students reported a higher total at posttest (W = 0, z = 4.02, p < .001, r = .88); mean normalized gain was 0.81 (high), with 14 high and seven medium individual gains. The dimension means increased for magnitude (2.74 to 3.79), strength (3.00 to 3.67), and generality (2.79 to 3.66), with significant paired differences in the reported tests. Two students had unchanged strength scores, and the item on explaining solutions publicly remained relatively

low at posttest (M = 3.57). These outcomes describe students' reported beliefs rather than direct observations of confidence in action.

Table 13. Mathematical self-efficacy before and after the intervention (n = 21)

Measure	Before M (SD)	After M (SD)	Gain	Wilcoxon z (r)
Total score (16–64)	45.43 (10.45)	59.24 (4.05)	+13.81	4.02 (.88)***
Magnitude (item mean)	2.74 (0.69)	3.79 (0.29)	+1.05	4.02 (.88)***
Strength (item mean)	3.00 (0.70)	3.67 (0.33)	+0.67	3.83 (.84)***
Generality (item mean)	2.79 (0.71)	3.66 (0.30)	+0.87	4.03 (.88)***
Mean N-gain (category)			0.81 (high)	

Each self-efficacy dimension increased in the reported paired analyses, with magnitude showing the largest mean change. Interpretation remains limited to self-reported efficacy beliefs. Figure 7 supplements the total scores with dimension profiles and N-gain distributions.

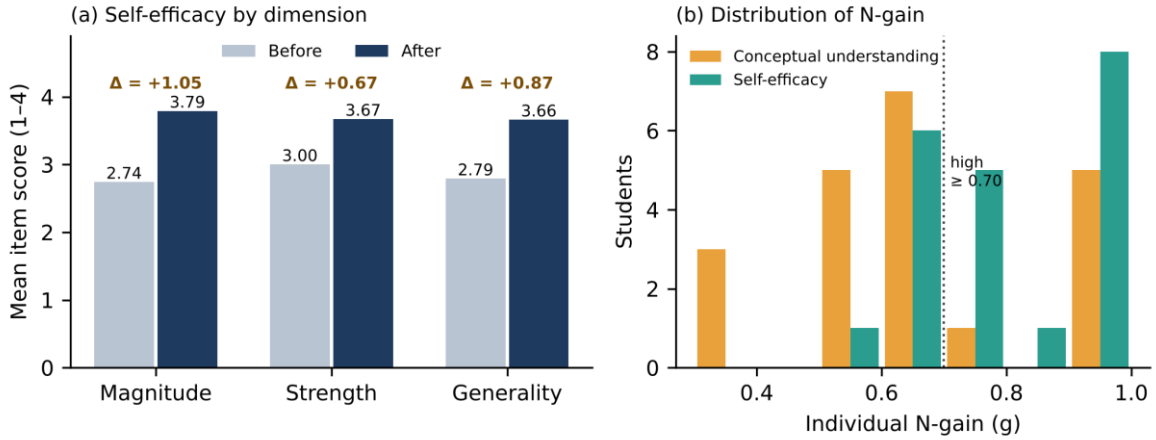


Figure 7. (a) Self-efficacy by dimension before and after the intervention and (b) distribution of individual N-gain for both outcomes

Magnitude displayed the largest self-efficacy mean change, whereas the conceptual-understanding gains were predominantly medium. The two scales should not be treated as interchangeable measures of learning. Exploratory Spearman analyses yielded a positive association between conceptual understanding and self-efficacy at pretest ($\rho = .47$, $p = .030$) and at posttest ($\rho = .62$, $p = .003$), whereas the individual gain scores showed little association ($\rho = .02$, $p = .914$). The two cross-sectional correlations should not be read as evidence that the relationship itself changed significantly, because that difference was not tested. Figure 7 summarizes the dimension-level self-efficacy profiles and the distributions of individual gains for the two measures.

D. Disseminate stage

Dissemination was deliberately limited. The final version of the site was shared with the mathematics teacher of the trial school before it was used with students, so that she could review every component and guide students during the lessons. The site remains openly accessible to the teacher for further use with other classes. Wider dissemination to other teachers and schools was beyond the scope of this study and is recommended as the next step.

Discussion

The development process produced a website whose content, media, and practitioner ratings satisfied the study's validity thresholds. Its relatively high content and practitioner assessments are consistent with a prototype developed from observable classroom needs and then organized through a phase-by-phase design matrix. Similar Google Sites studies have reported favorable expert assessments, but such convergence primarily concerns judged suitability rather than evidence of learning causation (Arnisya et al., 2025; Kristanto & Afifah, 2022; Rahmafani & Sulistyaningrum, 2025). The media expert's lower rating, especially for graphics and ease of use, identified a meaningful limitation of the initial prototype. Captions, textual instructions, and corrected prompts were consequently added, yet no second media rating was reported, so improvement in validation scores cannot be inferred from the revision itself. This distinction matters because international research on digital mathematics education attributes instructional value to designed interactions and representational connections, rather than to interface features alone (Bach et al., 2024; Engelbrecht & Borba, 2024).

Practicality evidence suggests that the integrated environment could be implemented in the participating class, although the sources of evidence were not equally independent. The implementation score of 92.22% and the student response score of 89.68% indicate that planned PBL activities were largely enacted and that students regarded the materials favorably. The 100% teacher-response score is less decisive because one teacher was both a respondent and an observer, and she had already reviewed the product. Ratings for the first meeting were also lower than for later meetings, which is compatible with students becoming familiar with the site, although no formal adaptation mechanism was tested. Studies of blended mathematical discussion and problem-based self-regulation likewise caution that navigation, communication roles, and learner independence must be actively supported (Giberti et al., 2022; Tempelaar et al., 2024). For this setting, the practical contribution lies in consolidating familiar tools into a sequence that the teacher could follow, rather than establishing superiority over other platforms.

Conceptual understanding improved during the six PBL meetings, with a mean increase of 21.29 points and a medium mean N-gain of 0.67. The result resembles favorable reports from Google Sites-supported mathematics instruction (Johdi et al., 2024; Mustika et al., 2024; Pakih et al., 2025; Riyanto et al., 2025). The design matrix provides a plausible instructional interpretation: contextual opening problems called for explanation before rule use, and LKPD investigations asked learners to move between examples, coordinates, and procedures. Research on dynamic representations and geometry likewise shows why the transition between visible objects and symbolic statements is a central learning demand (Bach et al., 2024; Gurmu et al., 2024; Valori et al., 2024). Nevertheless, because there was no comparison class or separate measure of exposure to each component, the data cannot isolate whether the observed change arose from PBL, Google Sites, the teacher's instruction, repeated testing, or their combination. The study accordingly supports feasibility and a promising pre–post pattern rather than a strong comparative effectiveness claim.

The contrast between complete posttest mastery and a medium N-gain deserves interpretation rather than being treated as a contradiction. Students crossed the local threshold of 75, yet improvement was constrained by both the difficulty of representations and the ceiling of the 0–100 assessment scale. Observer notes continued to identify problems with rotation direction, non-origin centers of dilation, and composition order, all of which require reasoning about transformed objects rather than reproducing a single coordinate rule. Research on representation quality suggests that connecting alternative displays explicitly can be more educationally consequential than multiplying illustrations (Dreher et al., 2024; Jäder & Johansson, 2025; Sánchez et al., 2024). The smaller dispersion in posttest scores suggests a more similar level

of performance across the class, but does not identify a causal mechanism or prove that students with the lowest pretest scores gained most. An item-by-indicator analysis would be needed to determine whether conceptual processes improved unevenly across the six indicators.

Reported mathematical self-efficacy increased substantially, particularly in magnitude, the belief that one can handle tasks of increasing difficulty. This direction is compatible with Bandura's (1997) account of mastery experiences and with evidence that structured mathematical tasks, confidence support, and strategy-focused feedback can shape task-specific beliefs (Herset et al., 2023; Koponen et al., 2021; Zakariya, 2022). Within the present intervention, graded worksheets created opportunities for repeated problem solving, while group presentations and teacher feedback could have provided models and encouragement. Similar research has examined the relationship between confidence, mathematical performance, and engagement, sometimes with different patterns across tasks or learners (Holenstein et al., 2022; Krawitz et al., 2026; Vogelsanger-Holenstein et al., 2025). These studies make the proposed explanation reasonable, but they do not establish that mastery experience was the measured mediator in this trial. Self-report increases should also be interpreted with possible novelty and social-desirability responses in mind.

The lower increase in strength compared with magnitude adds specificity to the affective findings. Confidence that extends to difficult tasks may be more responsive to short sequences of successful solutions than belief persistence when students encounter setbacks. Two students showed no change in strength, while the relative difficulty of explaining solutions to the whole class persisted in questionnaire results and observer accounts. Research on mathematical anxiety, metacognitive regulation, and motivational profiles supports treating confidence as situated and intertwined with task conditions rather than a fixed trait (Broda et al., 2023; Demedts et al., 2022; Hernández de la Hera et al., 2023; Říčan et al., 2022; Shimizu, 2025). The next iteration could therefore allocate more time to rehearsed small-group explanations and formative reflection, although these are design implications, not tested improvements. The different gains in conceptual understanding and self-efficacy should not be ranked as directly comparable magnitudes because the constructs and scoring scales differ.

The relationship between the cognitive and affective outcomes was exploratory and less straightforward than the mean gains alone might suggest. Both pretest and posttest correlations were positive, agreeing broadly with literature that associates mathematical self-efficacy with achievement and learning-related behavior (Sakellariou, 2022; Street et al., 2024; Wang et al., 2022; Yu et al., 2023). Yet individual gain scores were essentially unrelated ($\rho = .02$), indicating that students who changed more on one measure were not consistently those who changed more on the other. This is compatible with the possibility of partially distinct processes, such as developing representations through practice and confidence through perceived success, but the study did not test those explanations. Work examining task interpretation, feedback, and self-regulation similarly discourages assuming that cognitive and motivational changes always move together (Gentner et al., 2024; Navarro et al., 2025; Zakariya, 2021). The descriptive difference between the two time-specific correlations does not demonstrate a statistically significant strengthening of their association.

The principal limitation is the one-group pretest–posttest design, which cannot separate the website intervention from maturation, testing effects, or ordinary instruction. Purposive selection of 21 students from one advanced mathematics class limits population inference, and one validator per expertise area restricts the breadth of product-validity evidence. Implementation ratings were produced by a single observer, self-efficacy was measured only by questionnaire, and wider dissemination was not undertaken. Further studies could compare intact classes, include delayed measures, use multiple validators, analyze the six conceptual indicators individually, and triangulate self-report with classroom behavior. International studies

of technology-supported problem solving and digital metacognition similarly emphasize the instructional conditions and evaluation methods that shape reported outcomes (Mannan et al., 2026; Vessonen et al., 2025). Despite these limitations, the contribution of this development study is a transferable planning logic: connect PBL phases, website affordances, and intended cognitive and affective targets before evaluating a prototype in classroom use.

CONCLUSION

The 4D development process yielded a PBL-based Google Sites environment for geometric transformations whose validity and practicality ratings were 88.61% and 93.97%, respectively. The design matrix connected the five PBL phases to site components, intended sources of self-efficacy, and indicators of conceptual understanding. During the field trial with 21 Grade XI students, conceptual understanding increased from 66.52 to 87.81 (mean N-gain = 0.67), and self-efficacy increased from 45.43 to 59.24 (mean N-gain = 0.81); paired differences on both measures were statistically significant. These findings support the product's suitability for the local classroom and document promising cognitive and affective changes during implementation. They do not establish a causal advantage over alternative instruction because no comparison group was used. Future trials with comparison classes, repeated validation, and longer follow-up can establish whether the observed patterns extend beyond this school and which design features matter most.

ACKNOWLEDGMENTS

The authors thank the principal, the mathematics teacher, and the Grade XI students of the participating school, as well as the validators who reviewed the product and instruments.

DECLARATIONS

Author contributions. Risna Helmi: conceptualisation, investigation, product development, data curation, formal analysis, writing – original draft. Yumiati: supervision, methodology, validation, writing – review and editing. Endang Wahyuningrum: supervision, validation of instruments, writing – review and editing.

Conflict of interest. The authors declare no competing interests. None of the expert validators of the product is an author of this manuscript.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval and informed consent. Research permission was granted by the Graduate School of Universitas Terbuka and by the principal of the participating school. The intervention was part of regular instruction, students were informed of the purpose of the study, participation in the questionnaires was voluntary, and all data were anonymised.

Data availability. The anonymised datasets are available from the corresponding author upon reasonable request.

Use of generative AI. The authors used an AI-assisted tool to support language editing, text structuring, and verification of statistical calculations. The authors reviewed and edited all content and take full responsibility for the publication.

REFERENCES

Aba-Oli, Z., Koyas, K., & Husen, A. (2025). Higher-order thinking skills-oriented problem-based learning interventions in mathematics: A systematic literature review. *School Science and Mathematics, 125*(3), 214–231. https://doi.org/10.1111/ssm.12676

Afifah, S., Tamrin, M., Salsabila, K. I., Hasanah, A., & Herman, T. (2024). Analisis kemampuan siswa pada pemahaman konsep matematis materi barisan dan deret. *Jurnal Jendela Matematika, 2*(1), 11–20.

Arends, R. I. (2012). *Learning to teach* (9th ed.). McGraw-Hill.

Arnisya, M. F., Luritawaty, I. P., & Sumartini, T. S. (2025). Pengembangan media pembelajaran matematika berbantuan Google Sites pada materi sistem persamaan linear dua variabel. *Jurnal Pendidikan Matematika Universitas Lampung, 12*(4), 293–304. https://doi.org/10.23960/mtk/v12i4.pp293-304

Bach, C. C., Bergqvist, E., & Jankvist, U. T. (2024). Students' dynamic communication while transforming mathematical representations in a dynamic geometry environment. *ZDM – Mathematics Education, 56*, 543–557. https://doi.org/10.1007/s11858-024-01575-x

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Broda, M. D., Ross, E., Sorhagen, N., & Ekholm, E. (2023). Exploring control-value motivational profiles of mathematics anxiety, self-concept, and interest in adolescents. *Frontiers in Psychology, 14*, Article 1140924. https://doi.org/10.3389/fpsyg.2023.1140924

Cevikbas, M., & Kaiser, G. (2023). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. *ZDM – Mathematics Education, 55*, 177–191. https://doi.org/10.1007/s11858-022-01388-w

Ciung, M. V., Istiqomah, I., & Taufiq, I. (2022). Pengembangan media pembelajaran matematika berbasis Google Sites pada materi deret aritmatika. *Circle: Jurnal Pendidikan Matematika, 2*(1), 41–50. https://e-journal.uingusdur.ac.id/circle/article/view/332

Demedts, F., Reynvoet, B., Sasanguie, D., & Depaepe, F. (2022). Unraveling the role of math anxiety in students' math performance. *Frontiers in Psychology, 13*, Article 979113. https://doi.org/10.3389/fpsyg.2022.979113

Dreher, A., Wang, T.-Y., Feltes, P., Hsieh, F.-J., & Lindmeier, A. (2024). High-quality use of representations in the mathematics classroom: A matter of the cultural perspective?. *ZDM – Mathematics Education, 56*, 965–980. https://doi.org/10.1007/s11858-024-01597-5

Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM – Mathematics Education, 56*, 281–292. https://doi.org/10.1007/s11858-023-01530-2

Fitrianingsih, N., Apriliana, S., Ilmi, A. K., & Fakhriyana, D. (2024). Systematic literature review: Self-efficacy terhadap kemampuan pemahaman konsep matematis. *El Banar: Jurnal Pendidikan dan Pengajaran, 7*(2), 1–14. https://doi.org/10.54125/elbanar.v7i02.260

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2–18. https://doi.org/10.1037/a0024338

Gentner, N. M., Respondek, L., & Seufert, T. (2024). Effects of short- and long-term prompting in learning journals on strategy use, self-efficacy, and learning outcomes. *Instructional Science, 52*, 919–950. https://doi.org/10.1007/s11251-024-09671-x

Giberti, C., Arzarello, F., Bolondi, G., & Demo, H. (2022). Exploring students' mathematical discussions in a multi-level hybrid learning environment. *ZDM – Mathematics Education, 54*, 403–418. https://doi.org/10.1007/s11858-022-01364-4

Gurmu, F., Tuge, C., & Hunde, A. B. (2024). Effects of GeoGebra-assisted instructional methods on students' conceptual understanding of geometry. *Cogent Education, 11*(1), Article 2379745. https://doi.org/10.1080/2331186X.2024.2379745

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

Hardita, F., Zakiah, N. E., & Solihah, S. (2024). Kemampuan pemahaman konsep matematis siswa pada materi persegi dan persegi panjang. *Proceeding GAMMA-NC, 4*(1), 128–138. https://jurnal.unigal.ac.id/GAMMA-NC/article/download/15072/pdf

Hartmann, L.-M., Krawitz, J., & Schukajlow, S. (2021). Create your own problem! When given descriptions of real-world situations, do students pose and solve modelling problems?. *ZDM – Mathematics Education, 53*, 919–935. https://doi.org/10.1007/s11858-021-01224-7

Hernández de la Hera, J. M., Morales-Rodríguez, F. M., Rodríguez-Gobiet, J. P., & Martínez-Ramón, J. P. (2023). Attitudes toward mathematics/statistics, anxiety, self-efficacy and academic performance: An artificial neural network. *Frontiers in Psychology, 14*, Article 1214892. https://doi.org/10.3389/fpsyg.2023.1214892

Herset, M., El Ghami, M., & Bjerke, A. H. (2023). The effect of level-marked mathematics tasks on students' self-efficacy: An experimental study. *Frontiers in Psychology, 14*, Article 1116386. https://doi.org/10.3389/fpsyg.2023.1116386

Hidayat, A., & Firmanti, P. (2024). Navigating the tech frontier: A systematic review of technology integration in mathematics education. *Cogent Education, 11*(1), Article 2373559. https://doi.org/10.1080/2331186X.2024.2373559

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn?. *Educational Psychology Review, 16*(3), 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Holenstein, M., Bruckmaier, G., & Grob, A. (2022). How do self-efficacy and self-concept impact mathematical achievement? The case of mathematical modeling. *British Journal of Educational Psychology, 92*(1), 155–174. https://doi.org/10.1111/bjep.12443

Jäder, J., & Johansson, H. (2025). Exploring students' conceptual understanding through mathematical problem solving: Students' use of and shift between different representations of rational numbers. *Research in Mathematics Education.* https://doi.org/10.1080/14794802.2025.2456840

Jannah, M., Apriandi, D., & Andari, T. (2024). Pengembangan media pembelajaran berbasis Google Site interaktif untuk meningkatkan adversity quotient matematis siswa. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 13*(1), 293–303. https://doi.org/10.24127/ajpm.v13i1.8330

Johdi, H., Gunawan, G., Ayub, S., & Kosim, K. (2024). The effectiveness of interactive Google Sites-based learning media on students' conceptual understanding. *Indonesian Journal of STEM Education, 6*(2), 55–62. https://journal.publication-center.com/index.php/ijse/article/view/1638

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press. https://doi.org/10.17226/9822

Koponen, T., Aro, T., Peura, P., Leskinen, M., Viholainen, H., & Aro, M. (2021). Benefits of integrating an explicit self-efficacy intervention with calculation strategy training for low-performing elementary students. *Frontiers in Psychology, 12*, Article 714379. https://doi.org/10.3389/fpsyg.2021.714379

Krawitz, J., Meyer-Jenßen, L., Krausmüller, K., & Rakoczy, K. (2026). Problem posing and motivation: The effects of posing and solving one's own modelling problems on autonomy, competence, relatedness, and self-efficacy. *ZDM – Mathematics Education, 58*, 621–636. https://doi.org/10.1007/s11858-025-01762-4

Kristanto, I., & Afifah, D. S. N. (2022). SIJAMET (Sijago Matematika): Media pembelajaran berbasis Google Sites pada materi peluang. *JKPM (Jurnal Kajian Pendidikan Matematika), 8*(1), 63–74. https://doi.org/10.30998/jkpm.v8i1.14076

Lenz, K., Reinhold, F., & Wittmann, G. (2024). Transitions between conceptual and procedural knowledge profiles: Patterns in understanding fractions and indicators for individual differences. *Learning and Individual Differences, 116*, Article 102548. https://doi.org/10.1016/j.lindif.2024.102548

Mannan, A. M. K., Bajuri, M. R., Abdul Rahim, S. S., & Abdullah, N. A. (2026). How do digital tools support metacognition in mathematics learning? A systematic review of research on GeoGebra and problem-solving. *Educational Research Review, 52*, Article 100827. https://doi.org/10.1016/j.edurev.2026.100827

Mustika, G. S., Astuti, D., Aprilia, A., & Yuliandari, Y. (2024). Implementasi model problem-based learning berbantuan Google Sites untuk meningkatkan kemampuan pemahaman konsep siswa pada materi fungsi kuadrat. *Jurnal Inovasi Pembelajaran Matematika, 3*(1), 23–39. https://journal.rafandhapress.com/jipm/article/view/98

Navarro, M., Larrain, M., & Pezoa, J. (2025). Internal structure of a scale of sources of self-efficacy in mathematics among middle-school students. *International Journal of Educational Research, 134*, Article 102771. https://doi.org/10.1016/j.ijer.2025.102771

Negara, F. P., Abidin, Z., & Faradiba, S. S. (2023). Meningkatkan self-efficacy matematika siswa melalui pembelajaran berbasis masalah. *Jurnal Cendekia: Jurnal Pendidikan Matematika, 7*(1), 455–466. https://doi.org/10.31004/cendekia.v7i1.1943

Noverlika, R., Mujahidawati, M., & Falani, I. (2024). Pengembangan media pembelajaran Wegos (Web Google Sites) berbasis problem-based learning untuk meningkatkan kemampuan pemecahan masalah matematis siswa. *Jurnal Pendidikan MIPA, 14*(2), 365–372. https://doi.org/10.37630/jpm.v14i2.1524

Nst, M. B., Surya, E., & Khairani, N. (2023). Pengaruh model problem-based learning terhadap kemampuan pemecahan masalah matematika dan self-efficacy siswa. *Jurnal Cendekia: Jurnal Pendidikan Matematika, 7*(2), 1533–1544. https://doi.org/10.31004/cendekia.v7i2.2291

Pakih, S. M., Sugandi, A. I., & Zanthy, L. S. (2025). Penerapan model pembelajaran problem-based learning berbantuan Google Sites untuk meningkatkan kemampuan pemahaman matematis siswa SMP. *JPMI: Jurnal Pembelajaran Matematika Inovatif, 8*(5), 583–598. https://journal.ikipsiliwangi.ac.id/jpmi/article/view/28938

Rahmafani, E. A., & Sulistyaningrum, H. (2025). Pengembangan media pembelajaran interaktif berbantuan Google Sites untuk meningkatkan pemahaman konsep siswa SMP. *Jurnal Cendekia: Jurnal Pendidikan Matematika, 9*(3), 1238–1249. https://doi.org/10.31004/cendekia.v9i3.4363

Říčan, J., Chytrý, V., & Medová, J. (2022). Aspects of self-regulated learning and their influence on the mathematics achievement of fifth graders in the context of four different proclaimed curricula. *Frontiers in Psychology, 13*, Article 963151. https://doi.org/10.3389/fpsyg.2022.963151

Riyanto, H. W., Degeng, I. N. S., & Sitompul, N. C. (2025). The effect of Google Sites on elementary students' conceptual understanding in mathematics. *Jurnal Teknologi Pendidikan: Jurnal Penelitian dan Pengembangan Pembelajaran, 10*(4), 840–850. https://doi.org/10.33394/jtp.v10i4.18212

Sakellariou, C. (2022). The reciprocal relationship between mathematics self-efficacy and mathematics performance in US high school students: Instrumental variables estimates and gender differences. *Frontiers in Psychology, 13*, Article 941253. https://doi.org/10.3389/fpsyg.2022.941253

Sánchez, E., García-Ríos, V. N., & Sepúlveda, F. (2024). Development of high school students' conceptions of sampling distribution in the context of learning significance tests with technology. *Educational Studies in Mathematics, 117*, 215–238. https://doi.org/10.1007/s10649-024-10330-8

Shimizu, Y. (2025). Relation between mathematics self-efficacy, mathematics anxiety, behavioural engagement, and mathematics achievement in Japan. *Psychology International, 7*(2), Article 36. https://doi.org/10.3390/psycholint7020036

Sihombing, B., Zamsiswaya, & Sawaluddin. (2024). Model pengembangan 4D (Define, Design, Develop, dan Disseminate) dalam pembelajaran pendidikan Islam. *Journal of Islamic Education El Madani, 4*(1), 11–19. https://doi.org/10.55438/jiee.v4i1.135

Stovner, R. B., & Klette, K. (2022). Teacher feedback on procedural skills, conceptual understanding, and mathematical practices: A video study in lower secondary mathematics classrooms. *Teaching and Teacher Education, 110*, Article 103593. https://doi.org/10.1016/j.tate.2021.103593

Street, K. E. S., Malmberg, L.-E., & Schukajlow, S. (2024). Students' mathematics self-efficacy: A scoping review. *ZDM – Mathematics Education, 56*, 265–280. https://doi.org/10.1007/s11858-024-01548-0

Sudianto, S., & Ramdiani, R. (2024). SLR: Pengaruh model PBL terhadap kemampuan matematis dan sikap siswa. *Polinomial, 3*(1), 36–44. https://doi.org/10.56916/jp.v3i1.864

Tempelaar, D., Bátori, A., & Giesbers, B. (2024). Understanding self-regulation strategies in problem-based learning through dispositional learning analytics. *Frontiers in Education, 9*, Article 1382771. https://doi.org/10.3389/feduc.2024.1382771

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Leadership Training Institute/Special Education, University of Minnesota. https://files.eric.ed.gov/fulltext/ED090725.pdf

Uegatani, Y., Otani, H., Shirakawa, S., & Ito, R. (2024). Real and illusionary difficulties in conceptual learning in mathematics: Comparison between constructivist and inferentialist perspectives. *Mathematics Education Research Journal, 36*, 895–915. https://doi.org/10.1007/s13394-023-00478-6

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796. https://doi.org/10.3102/0034654308321456

Valori, G., Albanese, V., Adamuz-Povedano, N., & Gómez-Torres, E. (2024). Effectiveness of paper folding and GeoGebra on students' geometric thinking and visualization. *International Journal of Education in Mathematics, Science and Technology, 12*(5), 1395–1414. https://eric.ed.gov/?id=EJ1452836

Vessonen, T., Hellstrand, H., Kurkela, M., Aunio, P., & Laine, A. (2025). The effectiveness of mathematical word problem-solving interventions among elementary schoolers: A systematic review and meta-analysis. *International Journal of Educational Research, 132*, Article 102642. https://doi.org/10.1016/j.ijer.2025.102642

Vogelsanger-Holenstein, M., Schukajlow, S., & Bruckmaier, G. (2025). Mathematical modelling and self-efficacy: Immediate and long-lasting effects of teaching mathematical modelling with a solution plan. *ZDM – Mathematics Education, 57*, 489–502. https://doi.org/10.1007/s11858-025-01674-3

Wang, L., Peng, F., & Song, N. (2022). The impact of students' mathematical attitudes on intentions, behavioral engagement, and mathematical performance in the China's context. *Frontiers in Psychology, 13*, Article 1037853. https://doi.org/10.3389/fpsyg.2022.1037853

Waseso, K., Fitriasari, P., & Isroqmi, A. (2022). Inovasi media pembelajaran matematika interaktif berbasis Google Sites. *AKSIOMA: Jurnal Matematika dan Pendidikan Matematika, 13*(2), 301–314. https://doi.org/10.26877/aks.v13i2.12018

Yang, Z., Yang, X., Wang, K., Zhang, Y., Pei, G., & Xu, B. (2021). The emergence of mathematical understanding: Connecting to the closest superordinate and convertible concepts. *Frontiers in Psychology, 12*, Article 525493. https://doi.org/10.3389/fpsyg.2021.525493

Yu, W., Zhou, S., & Zhou, Y. (2023). Measuring mathematics self-efficacy: Multitrait-multimethod comparison. *Frontiers in Psychology, 14*, Article 1108536. https://doi.org/10.3389/fpsyg.2023.1108536

Zakariya, Y. F. (2021). Self-efficacy between previous and current mathematics performance of undergraduate students: An instrumental variable approach to exposing a causal relationship. *Frontiers in Psychology, 11*, Article 556607. https://doi.org/10.3389/fpsyg.2020.556607

Zakariya, Y. F. (2022). Improving students' mathematics self-efficacy: A systematic review of intervention studies. *Frontiers in Psychology, 13*, Article 986622. https://doi.org/10.3389/fpsyg.2022.986622

Živković, M., Pellizzoni, S., Doz, E., Cuder, A., Mammarella, I., & Passolunghi, M. C. (2023). Math self-efficacy or anxiety? The role of emotional and motivational contribution in math performance. *Social Psychology of Education, 26*, 579–601. https://doi.org/10.1007/s11218-023-09760-8