Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

Tracing Students' Productive Struggle and Metacognitive Regulation through Written Solutions to Mathematical Problem Solving

Yhasinta Agustyarini[1)*], Dwi Juniati[2)], Siti Khabibah[3)]

Universitas KH. Abdul Chalim[1)]; Universitas Negeri Surabaya[2) 3)]

*Corresponding Email: yhasinta2018@gmail.com

Abstract. *This study follows students' struggle and looks for signs of metacognitive regulation while they solve written math problems. The study uses a qualitative design and looks at six photographs of handwritten solutions made by three students. Each student worked on a math task: unit-fraction equations (S1) rectangle geometry (S2) and sums of consecutive positive integers (S3). The written work was split into eleven segments called episodes. Each episode was checked for math obstacles changes in strategy, verification steps and how complete the justifications were. Planning, monitoring, evaluation and regulation were used as categories to interpret visible marks. They were not used to measure the students' thoughts directly. S1 moved from trying substitutions to counting within limits. Found all ten valid triples. S2 swapped a Pythagorean assumption for a justified extra construction and got correct numbers but a broad geometric identity was still not proved. S3 went from testing how long sequences could be to using factorization and parity arguments. S3 proved that 1024 could not be represented. Did not prove a wider guess, about powers of two. Across all cases productive struggle showed up as math progress after an obstacle. Metacognitive regulation could only be guessed from the written strategy changes. The findings show the usefulness and limits of written solutions when studying math problem-solving.*

Keywords: *productive struggle; metacognitive regulation; mathematical problem solving*

INTRODUCTION

Mathematical problem solving often requires students to make decisions when an initially plausible approach is insufficient. In this study, mathematical problem solving refers to working through mathematical tasks whose solution pathways are not immediately available to the solver. The three tasks examined are nonroutine examples of this activity. The title therefore highlights problem solving while the task descriptions still have a nonroutine nature. A student may find one number triple without showing that the solution is complete may apply a geometric theorem without checking its conditions or may test several integer sequences without proving a general claim. Such obstacles call

for mathematical reasoning about constraints, representations, and justification. Schoenfeld emphasized the importance of control decisions during problem solving (SCHOENFELD, 1985), and Santos-Trigo (2024) situated such decisions within the development of mathematical problem-solving research. The analysis below asks how these decision points are reflected in students' written work rather than assuming that the finished artifact reveals their entire cognitive experience.

Variation in how students approach mathematical problems has been documented across learner groups and school contexts (Amalina & Vidákovich, 2023). Related work has examined productive struggle alongside mathematical problem-solving skills among elementary school pupils (Mufliva et al., 2024). These studies provide a useful backdrop, but they do not justify treating a single written solution as a stable measure of a student's ability or disposition. In the present analysis, each case is interpreted in relation to the specific demands of its task.

Metacognition helps explain why two equally lengthy solutions do not necessarily reflect equally effective problem-solving processes. Flavell identified the monitoring and regulation of cognition as central components of metacognition (Flavell, 1979). In mathematical problem solving, these ideas are closely related to questions that may arise during students' work: What is already known? Does this calculation still satisfy the constraints? Does my method produce examples only, or can it establish that every possibility has been considered? When does a conjecture require proof? Research has linked mathematical metacognition with performance, particularly on tasks involving complex reasoning (Muncer et al., 2022); Xie et al., 2024). Yet associations between scores do not by themselves explain how an individual decision is made in the course of solving a particular problem.

Studies of metacognitive activity also vary in their participants and forms of evidence. Kholid and Ahadiyati examined students' metacognition in nonroutine problems (Kholid & Ahadiyati, 2022), while Kusaka and Ndihokubwayo studied metacognitive strategies in mathematical word problems (Kusaka & Ndihokubwayo, 2022). Further perspectives come from work with prospective teachers (Ramlah et al., 2024) and secondary students engaged in collaborative problem solving (Kusaka & Claude Habimana, 2025); (Prastiti et al., 2026). These studies help shape the way we talk

about the analysis even though information from talking and working together can't be used the same way, for pictures of individual written answers.

This is why traces of the solution process warrant close examination. Braithwaite and Sprague investigated the interplay of conceptual knowledge, procedural knowledge, and metacognition in routine and nonroutine problem solving (Braithwaite & Sprague, 2021). Baumanns and Rott proposed an operational framework for identifying planning, monitoring/control, and evaluation (Baumanns & Rott, 2023), while Zepeda and Nokes-Malach (2023) cautioned that different metacognitive measures do not necessarily capture the same processes. Written work can preserve valuable traces: a revised equation, a factorization reconsidered, a diagram supplemented with an auxiliary line, or a candidate answer rejected for violating a condition. Nevertheless, writing is not a recording of thought. Without interviews or recordings of the activity, a researcher cannot establish precisely when students hesitated, how long they paused, or which unrecorded personal reasons informed their decisions. Accordingly, the present study distinguishes what is visible on the page from plausible interpretations of those traces.

Metacognitive evidence also differs according to what an investigation measures. Oudman et al. explored self-scoring in relation to students' monitoring and regulation (Oudman et al., 2022). Aydın and Özgeldi considered metacognition in the relationship between test anxiety and mathematics achievement (Aydın & Özgeldi, 2024), while Scheibe et al. connected metacognitive cues, working memory, and mathematics anxiety in a theoretical framework (Scheibe et al., 2023). Such research highlights the need to identify what is—and is not—observable. The current worksheets do not establish students' anxiety, working-memory demands, or the accuracy of private self-judgments.

Recent studies characterize metacognitive regulation as context-sensitive and capable of changing during problem solving. Toikka et al. (2024) emphasized the importance of interpreting metacognitive knowledge in relation to the task, and Toikka et al. (2026) showed that planning, monitoring, and evaluation may not be outwardly apparent without opportunities for reflection. Torregrosa and Albarracín likewise highlighted reciprocal relationships between mathematical activity and self-regulation (Torregrosa & Albarracín, 2026). These findings lead to a question: when students notice a missing step what proof shows they really change how they act? In research that looks

at student work the answer has to come from changes that can be seen—like shifts, in symbols, diagrams or reasoning. It's not enough to guess what they were thinking. We need to see what they actually did.

The available evidence also differs across the settings in which mathematical regulation has been studied. These include floor-robot programming tasks (Del Olmo-Muñoz et al., 2024), mathematical problem solving with ChatGPT (Ihsani & Siswono, 2024), metacognitive instructional strategies for problem-solving competence (Rajadurai & Ganapathy, 2023), and practice on an online learning platform (Schonberg et al., 2026). The studies address different phenomena and rely on different records of activity. They are useful contrasts, rather than direct analogues, to a corpus consisting solely of photographed worksheets.

Productive struggle does not require students to remain confused or unsuccessful. Its significance lies in how difficulty creates an opportunity for improved mathematical understanding. Warshauer described several forms of struggle encountered in mathematical activity (Warshauer, 2015), while Granberg demonstrated that errors can become opportunities for learning when they are examined and addressed (Granberg, 2016). Crossing out an incorrect answer is therefore not, by itself, proof of a productive process. Stronger evidence is available when the correction is followed by an approach that enables mathematical progress. Biccard and Young et al. similarly emphasized opportunities for reasoning and the need to examine students' experiences at a finer level of detail (Biccard, 2024); (Young et al., 2024). This perspective is well suited to tasks that cannot be resolved through a single immediately apparent algorithm.

Research on support for productive struggle reminds teachers of the delicate balance between allowing time for independent thought and intervening to help (Stewart & Ball, 2023). Students' evaluations of the support they receive add another perspective to that discussion (Khalil et al., 2025). Before making a choice about what kind of help's best researchers need to understand what students work actually shows. A correct answer might come from a method that wasn't right and was changed later; on the hand an answer that isn't complete might still have useful math reasoning. This paper does not look at how many correct answers students get. Instead it looks at the steps they take—moving from trying numbers to setting limits from drawing pictures to writing equations and,

from seeing patterns to making strong arguments. Such transitions align with Schoenfeld's concept of mathematical control, which encompasses decisions to select, assess, and change strategies (SCHOENFELD, 1985).

One methodological difficulty is easily overlooked when studies reproduce excerpts of students' work: the analytical narrative can become more detailed than the underlying evidence. If a student writes "perhaps," for example, the researcher cannot turn that expression into a claim that the student has proved a theorem. Likewise, two approaches appearing on the same page do not establish how long an impasse lasted or the student's emotional sequence. At the same time, apparently minor annotations may be mathematically consequential. The note "38, not 36" indicates a discrepancy; the bound $x \leq 6$ turns an open-ended search into a finite one; and separating odd from even values of k makes it possible to prove that a representation cannot exist. Such traces offer concrete grounds for discussing productive struggle without inventing undocumented events.

Methodological work clarifies why this distinction matters. Verbal protocol analysis treats spoken reports as a particular kind of process evidence (Ericsson & Simon, 1993); Favier and Dorier proposed ways of analyzing heuristics and semantic spaces in students' mathematical work (Favier & Dorier, 2024); and Jin and Leong addressed how monitoring and regulation can be made visible in problem solving (Jin & Leong, 2025). None of these methods licenses the reconstruction of speech or unrecorded behavior from a finished worksheet. The present interpretations remain tied to visible inscriptions, mathematical validity, and the limits of each written claim.

The analytical framework brings productive struggle and metacognitive regulation into a carefully bounded relationship. Productive struggle is operationalized as an episode in which an identifiable mathematical obstacle is followed by continued engagement and a mathematically meaningful development, whether or not the final response is fully proved (Warshauer, 2015); (Granberg, 2016); (Young et al., 2024). Metacognitive regulation is considered through the functions that written actions may serve: planning a mathematical approach, monitoring an inconsistency, evaluating whether a result satisfies task constraints, and regulating the solution through a visible strategic adjustment (Flavell, 1979); (Baumanns & Rott, 2023). These labels describe evidence-based

interpretations of inscriptions rather than direct observations of private thought. In particular, a corrected equation alone does not establish a conscious monitoring episode, and a valid numerical pattern does not constitute a general proof.

The materials analyzed here consist of photographs of written solutions, not videos, spoken protocols, or interview transcripts. Their principal strength is that the symbolic work can be checked against mathematical rules. Their limitation is equally important: a photograph cannot reveal an unstated reason or an event that occurred outside the page. All process descriptions in the paper are therefore presented as artifact-based interpretations. The example interviews and think-aloud protocols in the appendix are explicitly labeled simulations and are provided only to illustrate how future research might test these interpretations. They are neither treated as participant quotations nor included as empirical research findings.

Accordingly, this study addresses three research questions: (1) What mathematical obstacles and episodes of productive struggle can be traced in students' written mathematical problem-solving solutions? (2) How do observable written actions provide tentative evidence of planning, monitoring, evaluation, and regulation as students adjust their solution strategies? (3) How does mathematical progress differ across the three cases in terms of valid results, rejected alternatives, and completeness of justification? The purpose is not to diagnose stable metacognitive abilities, but to explain the relationships among written evidence, strategic change, and mathematical validity in each task.

RESEARCH METHOD

This study employed an interpretive qualitative approach with three documentation-based cases of mathematical problem solving. Each case comprised one student's written solution to one nonroutine mathematical task. The unit of analysis was the written episode: a segment showing a mathematical objective, obstacle, revision, alternative representation, or check of the result. The analysis traced productive struggle through the development of mathematical arguments and cautiously considered the possible regulatory function of observable changes. Because the data consist of photographs of completed work rather than recordings of activity, the reconstructed sequence is limited

to defensible relationships among written traces. No unrecorded speech, emotions, intentions, or timing were treated as empirical findings.

The collection had three groups of answers each recorded in two pictures making a total of six pictures. S1 finished NR-1 a question about triples of numbers and a sum of three fractions that are one over a number; S2 finished NR-2, which was about a point inside a rectangle; and S3 finished NR-3, which was, about adding up consecutive positive numbers. The study included three students and three problems. Each student did not work on all three problems. The pictures do not show who the students were, when they worked how long they worked or how they were chosen; these details are not shared as confirmed research facts.

Table 1. Distribution of three students, three tasks, and six photographs of written artifacts.

Case	Task	Mathematical focus	Available photographs
S1 / NR-1	$\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=\frac{1}{2}$; $3 \le x \le y \le z$	Bound the search and prove that all triples of positive integers have been found	2 pages
S2 / NR-2	Rectangle with $AB = 6$, $PA^2 = 13$, $PB^2 = 25$, $PC^2 = 20$	Find BC, PD, the position of P, and a general relationship among four squared distances	2 pages
S3 / NR-3	Representing 2026 and 1024 as sums of consecutive positive integers	Find all representations, investigate 1024, and formulate and prove a general conjecture	2 pages

Sources of evidence and analytical procedure

The only primary materials available were photographs of handwritten solutions. Each complete page was examined before selecting excerpts showing a change in strategy or a judgment about the result. The content of excerpts reproduced in the findings was not altered; cropping and display adjustments served only to improve legibility. Short written comments were interpreted in their original sense, not transformed into imagined spoken dialogue. No audio recording, video of task completion, direct observational record, or interview was available as empirical evidence. Illustrative materials in the appendix are therefore kept strictly separate from the analytical corpus.

Analysis combined repeated reading with mathematical verification. First, each page was interpreted in relation to the task demand: finding every solution, determining dimensions and position, or proving whether a representation was possible. Next, written segments marking a change of direction were identified as episodes, such as establishing a bound on x, adding the auxiliary line PE, or distinguishing odd and even factors. Every episode was then checked against the task conditions to establish whether the mathematical step was valid and whether the conclusion was complete. Only then were the categories planning (PL), monitoring (MN), evaluation (EV), and regulation (RG) applied tentatively according to the function of the written action. Finally, cases were compared to identify types of strategic adjustment, not to rank the students' abilities.

Operational definitions for coding

Table 2. Coding metacognitive regulation from evidence in written work.

Code	Category	Indicator in the written artifact (not an attribution of mental state)
PL	Planning	Introducing variables; selecting a formula or diagram; setting subgoals for searching or proving.
MN	Monitoring	Noting discrepancies, crossing out work, checking results, or revisiting a condition when a provisional solution is problematic.
EV	Evaluation	Explicitly or implicitly judging whether a value is admissible, whether a strategy is sufficient, or whether the list of solutions is exhaustive.
RG	Regulation	A visible change after a difficulty: changing representation, narrowing the search, rejecting a root, revising a conjecture, or switching methods.

This episode seemed like a struggle. The student kept working on the math problem while the problem stayed unsolved. When the trace was later checked there was improvement showing that productive struggle had produced real progress.The qualifier "possible" is essential because photographs do not reveal effort intensity, pauses, or subjective experiences. An uncorrected arithmetic error, for instance, is insufficient evidence of regulation, whereas a conjecture not yet proved may still be preceded by productive reasoning. This distinction avoids equating every correction with

metacognition or judging productivity solely by the correctness of the final answer (Granberg, 2016); (Young et al., 2024).

Trustworthiness, ethics, and source limitations

Interpretive consistency was pursued by rereading all six photographs, independently checking the mathematical transformations, and recording distinctions between written facts and analytical inferences. Visible inaccuracies—including S3's incorrect verification sum—were retained in the analysis. The study did not have recordings for triangulation, independently coded judgments from two raters, or verifiable documentation of participant consent. Consequently, no intercoder reliability statistic is reported, and interviews are not presented as if they had taken place. Before this manuscript can serve as a final field-research report, the provenance of the artifacts, permission to use them, and the data-collection procedures must be verified.

RESULTS AND DISCUSSION

Case S1 (NR-1): From trial substitutions to bounding the complete solution set

S1's worksheet begins with simple numerical choices. The student writes $x = 3$ and $y = 3$, although these two unit fractions already add up to $\frac{2}{3}$. The next trial, $x = 3$ and $y = 6$, cannot work either: $\frac{1}{3} + \frac{1}{6}$ already equals $\frac{1}{2}$, even before the positive term $\frac{1}{z}$ is added. After increasing y to 7, S1 obtains $\frac{1}{z} = \frac{1}{42}$ and records the triple (3, 7, 42). These initial traces suggest an exploration grounded in concrete values. The first valid triple addresses part of the problem, but not the requirement to find "all triples." A meaningful shift becomes visible when the subsequent work goes beyond further substitution and uses the ordered fractions to bound x.

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

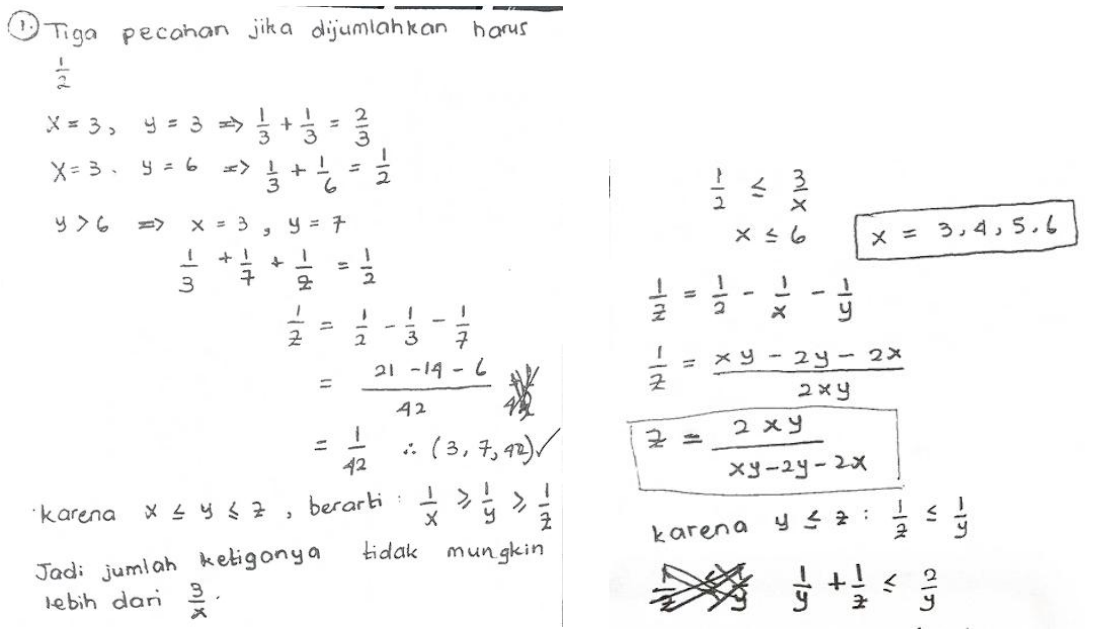
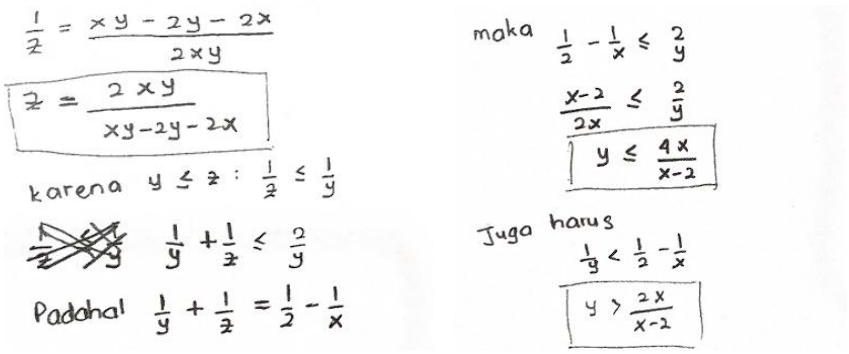


Figure 1. Excerpt from S1's handwritten work: trials with $x = 3$ and the bound $x \leq 6$ (cropped from an uploaded photograph of the original work).

S1 uses $x \leq y \leq z$ to write $\frac{1}{x} \geq \frac{1}{y} \geq \frac{1}{z}$. Since the three fractions sum to $\frac{1}{2}$, the inequality $\frac{1}{2} \leq \frac{3}{x}$ follows, giving $x \leq 6$. Along with $x \geq 3$, this leaves only four possible values: 3, 4, 5, and 6. The same sheet shows $z = \frac{2xy}{(xy - 2y - 2x)}$ and reasoning that leads to $y \leq \frac{4x}{(x-2)}$ and $y > \frac{2x}{(x-2)}$. Crossings-out near the inequality are worth noting, but they cannot be attributed to a particular feeling or personal reason without asking the student. The stronger evidence concerns the structure of the written work itself: an unbounded list of trials is replaced by a finite search based on mathematical constraints.



Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

Figure 2. Excerpt from S1: the formula for z and bounds on y, with visible revisions near the inequalities.

Having bounded the search, S1 examines the remaining cases. For $x = 3$, the values y = 7, 8, 9, 10, and 12 produce five valid triples; $y = 11$ is rejected because $z = \frac{66}{5}$ is not an integer. For $x = 4$, the student accepts (4, 5, 20), (4, 6, 12), and (4, 8, 8), while rejecting $z = \frac{28}{3}$. With $x = 5$, only (5, 5, 10) satisfies the conditions, and $x = 6$ yields (6, 6, 6). The final record contains ten triples, distributed as $5 + 3 + 1 + 1$. Mathematically, the solution is complete: x has been bounded, admissible values of y have been examined, and each resulting z has been checked for integrality and the required ordering. What matters most is not the number ten, but the justification for excluding every case outside this finite list.

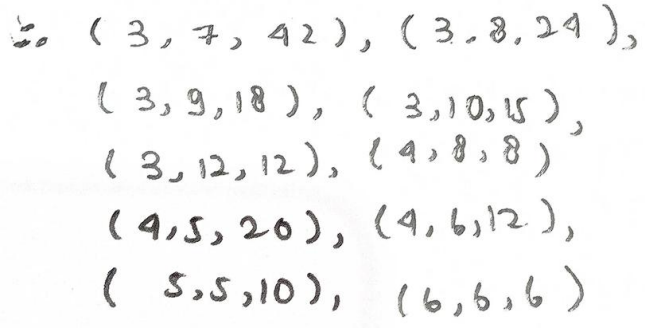


Figure 3. Excerpt from S1: the final list of ten triples, comprising all solutions under the stated constraints.

In S1, the written shift in strategy supports an interpretation of regulation oriented toward completeness. Bounding x and y controls the search space, while rejecting fractional values of z checks the task conditions. These actions differ from merely continuing to try values. Nonetheless, the photographs do not establish as facts that S1 "felt stuck" or "consciously recognized" a weakness in the initial approach. Productive struggle here refers to visible progress within the written solution, with explicit limits on claims about the student's internal experience.

Case S2 (NR-2): Correcting a Pythagorean assumption and reasoning with coordinates

S2 begins by drawing a rectangle and connecting P to several vertices. Below the sketch, the student writes the Pythagorean theorem and attempts to equate $PA^2 + PB^2$

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

with AB^2. The calculation $13 + 25 = 36$ is then followed by the handwritten observation "But 38 is not 36." This note indicates that the student has identified a numerical inconsistency. More fundamentally, triangle APB is not known to be right-angled, so the proposed relationship has no valid geometric basis. The photograph does not record S2's spoken explanation for abandoning it. However, the next construction provides observable evidence of change: a perpendicular is drawn from P to AB to create right triangles in which Pythagoras can legitimately be applied.

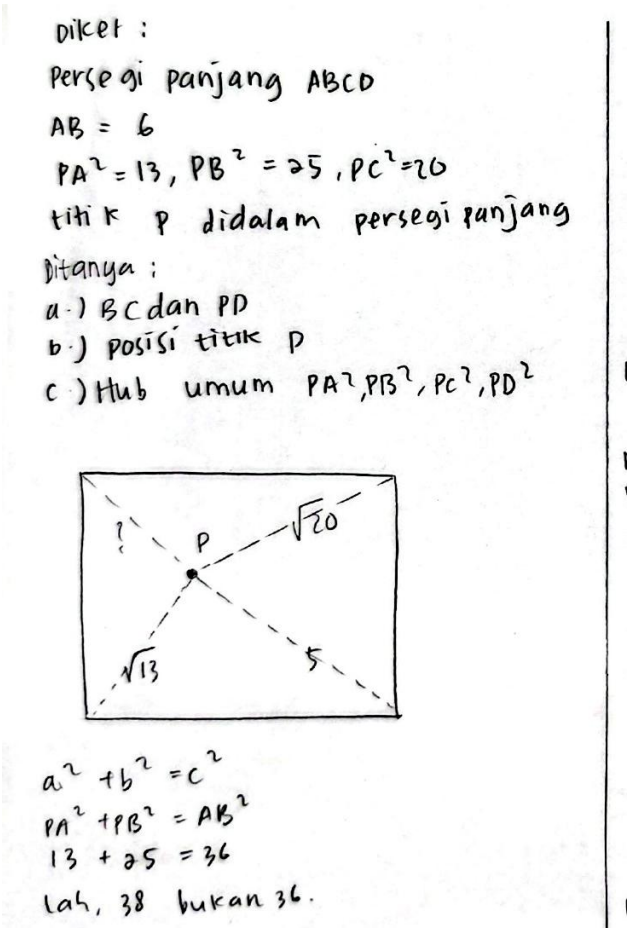


Figure 4. Excerpt from S2: an initial rectangle sketch and an attempted Pythagorean relationship marked as inconsistent.

The auxiliary line PE allows S2 to define $AE = x$, $PE = y$, and $EB = 6 - x$. The two right triangles yield $x^2 + y^2 = 13$ and $(6 - x)^2 + y^2 = 25$. Subtracting the equations eliminates y^2 and gives $x = 2$; substitution then yields $y^2 = 9$. S2 does not automatically accept both square roots. On the following photograph, the student explains in writing that $y = -3$ would place P below AB, contradicting the requirement

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

that P lie inside the rectangle. The admissible value is therefore $y = 3$. This segment presents two distinguishable checks: whether the geometric representation is valid and whether an algebraic candidate satisfies the geometric conditions.

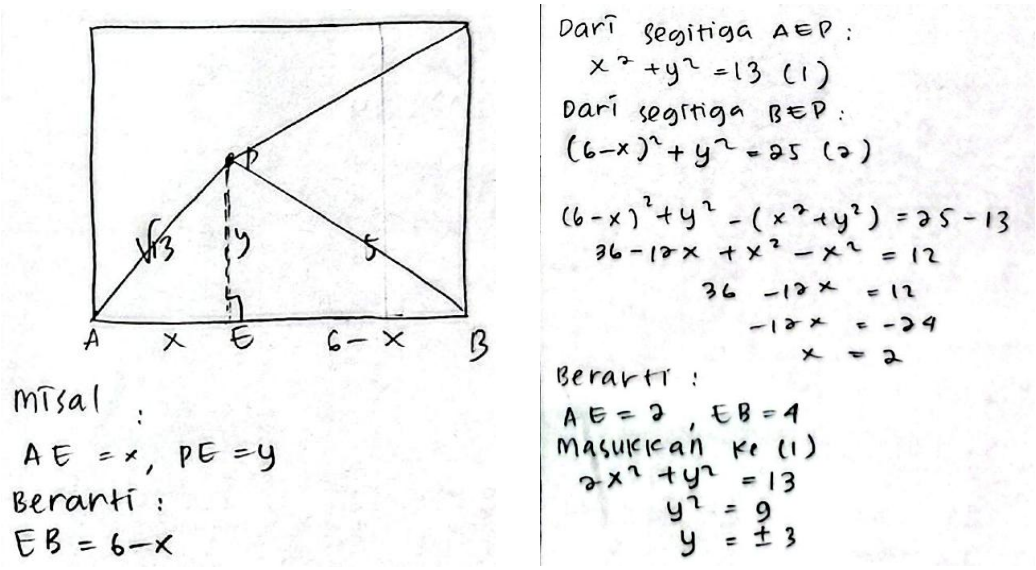


Figure 5. Excerpt from S2: the auxiliary line PE, the two equations, and the solutions $x = 2$ and $y = \pm 3$.

S2 then uses $PC^2 = 20$ to find the rectangle's height, writing $4^2 + (h - 3)^2 = 20$ and obtaining $h = 5$ or $h = 1$. The selected result, $BC = 5$, is consistent with P lying inside the rectangle. The written explanation for rejecting $h = 1$ is less explicit than the explanation for rejecting $y = -3$; the two episodes should therefore not be described as having identical evidential clarity. Once the height is determined, the student finds $PD^2 = 2^2 + 2^2 = 8$, hence $PD = 2\sqrt{2}$, and records $P = (2, 3)$. All these values are correct. The solution illustrates that checking geometric constraints is necessary at more than one point in the calculation.

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

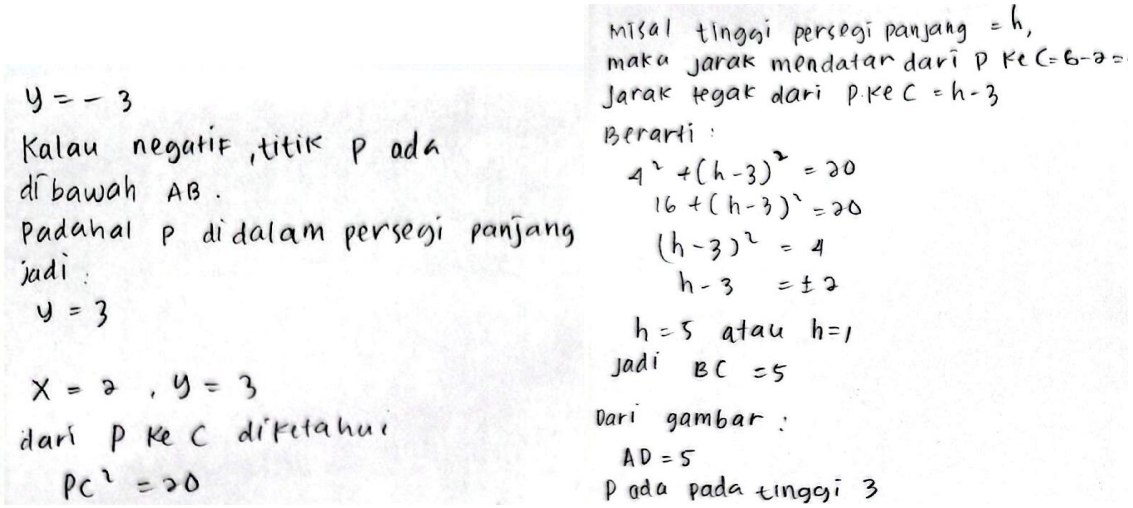


Figure 6. Excerpt from S2: rejection of $y = -3$ on positional grounds and the candidate heights $h = 5$ and $h = 1$.

Part (c) reveals a limitation that is particularly important for interpretation. S2 first adds $13 + 25$ and $20 + 8$; the totals are unequal. The student then pairs $13 + 20$ with $25 + 8$ and finds 33 in both cases. The handwritten phrase "So perhaps" precedes the boxed identity $PA^2 + PC^2 = PB^2 + PD^2$. Although this identity is true for every rectangle, the available work checks only one numerical configuration. No general coordinate derivation or algebraic proof appears in the photographs. Accordingly, S2's contribution to this part must be reported as a mathematically correct conjecture rather than a theorem the student has proved.

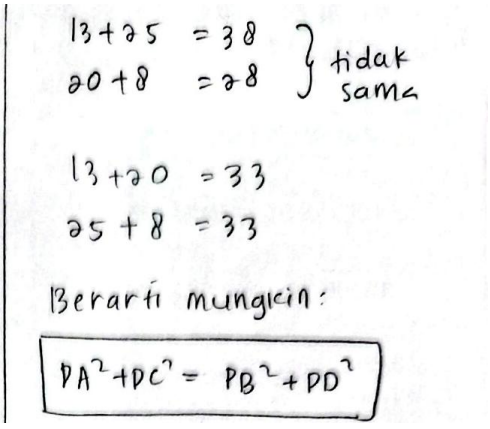


Figure 7. Excerpt from S2: numerical pattern-seeking and the conjecture $PA^2 + PC^2 = PB^2 + PD^2$; no general proof is visible.

S2's traces show that representation-based regulation is, in play. An unsuitable connection is replaced by a structure that supports a correct solution. Rejecting a

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

coordinate and selecting a practical height demonstrate that S2 pays attention to the limits.At the same time, the word "perhaps" in the final statement places a necessary limit on the researcher's claim. The student makes conceptual progress toward a general identity, but does not provide universal justification. This case illustrates how productive struggle may improve a strategy without satisfying every proof requirement of the task.

Case S3 (NR-3): From testing sequence lengths to reasoning about factors

In the task involving 2026, S3 begins by considering two numbers. The pair 1013 + 1013 is rejected because the numbers are equal, and 1012 + 1014 is rejected because their difference is two, not one. The student then models two consecutive integers as $a + (a + 1) = 2026$ and obtains $a = 1012.5$, which is not an integer. For three terms, S3 notes that 2026 is not divisible by three. Trying four terms finally gives $a = 505$ and the sequence 505 + 506 + 507 + 508. One mistake in the checking step must remain visible in the analysis: the student writes 1011 + 1015 = 2020, although the correct sum is 2026. This local arithmetic error does not invalidate the representation itself, but it indicates that the numerical verification was not entirely consistent.

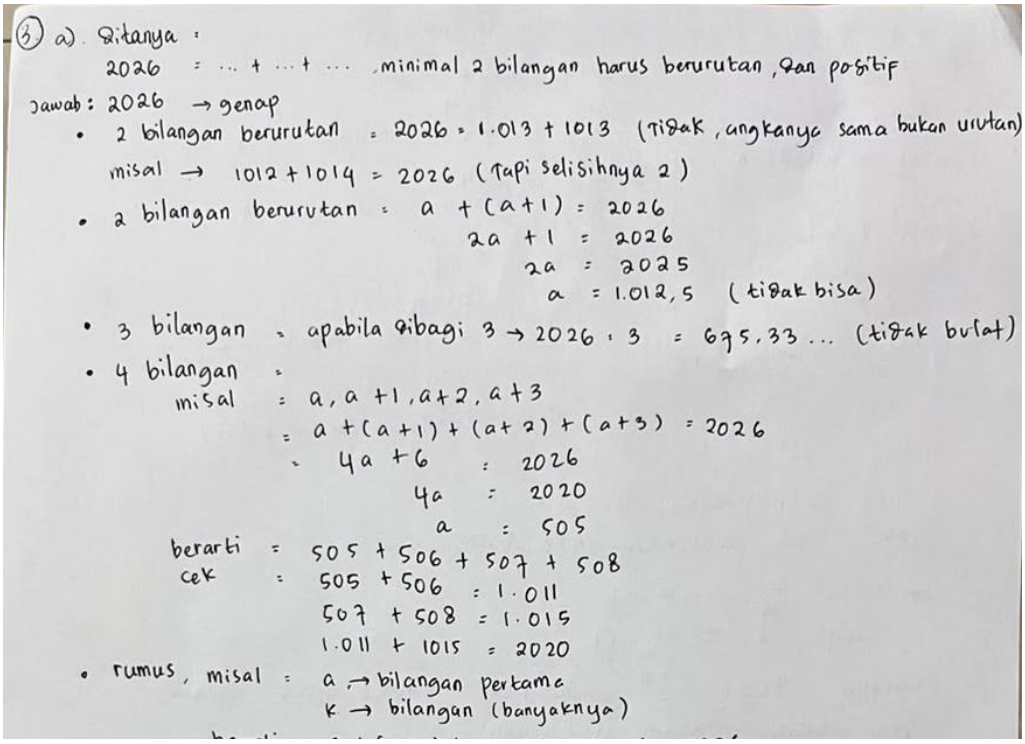


Figure 8. Excerpt from S3: trials of two to four terms, the representation 505 + 506 + 507 + 508, and the arithmetic slip in the checking line.

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

S3 does not assume that finding one representation proves the list complete. The student writes the general sum for k terms as $k(2a + k - 1) = 4052$. Factoring $4052 = 2^2 \times 1013$ provides possible values of k among its divisors. The smallest sum of k consecutive positive integers, $1 + 2 + ... + k$, then yields $k(k + 1) \leq 4052$ and therefore $k \leq 63$. Together, these conditions leave $k = 2$ or $k = 4$; the former has already failed, whereas the latter gives 505–508. A shift comparable to S1 is evident, but it operates differently: instead of bounding denominators, S3 bounds the number of terms. The mathematical question is no longer simply whether a representation exists, but why no other representation can exist.

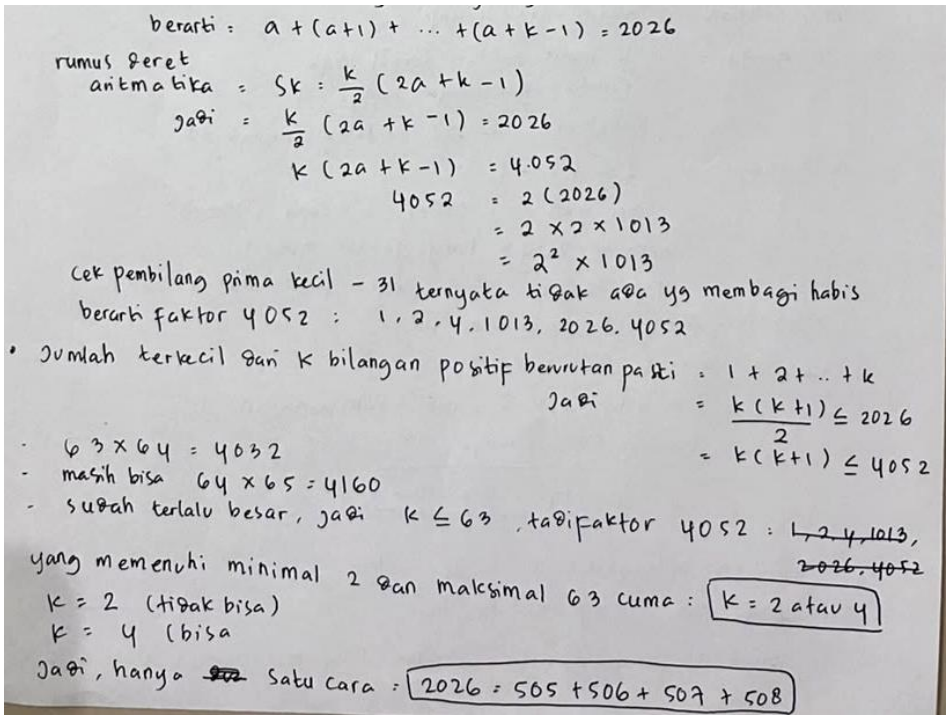


Figure 9. Excerpt from S3: the k-term sum, factorization of 4052, and the bound k ≤ 63 used to establish completeness.

Turning to 1024, S3 again tests sequences of two, three, four, and five terms. The student then writes $1024 = 2^{10}$ and reformulates the problem as $2048 = k(2a + k - 1)$. If k is odd, it must be an odd divisor of 2048; the only such divisor is 1, whereas $k \geq 2$. If k is even, the second factor is odd and, because $a \geq 1$ and $k \geq 2$, is at least 3. This is impossible because 2048 is a power of two. Unlike the failure of the four initial trials, the factorization argument covers every possible k and constitutes a valid proof that 1024 cannot be represented in the required form.

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

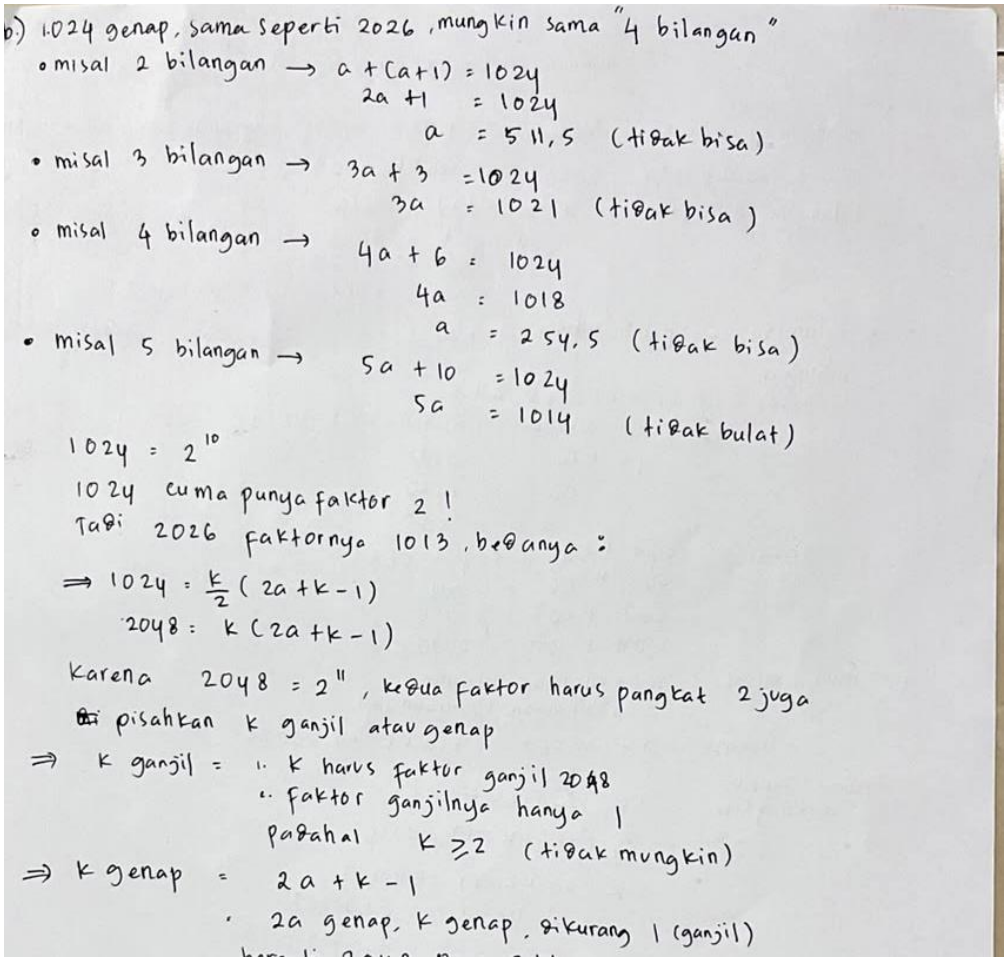


Figure 10. Excerpt from S3: testing short sequences for 1024 before developing a parity-of-factors argument.

After proving this special case, S3 reflects in writing that 1024 is not impossible merely because it is even: 2026 is also even and has a valid representation. The student writes, in effect, “Perhaps the reason is that 1024 is a power of two.” The tentative wording matters. In the artifact, this remains a reasoned conjecture, not a proved characterization of all positive integers. A proof for 1024 alone does not establish both that every power of two is impossible to represent and that every non-power of two is representable. Respecting this boundary allows S3’s case to illuminate the difference between a valid proof of one instance and a generalization that remains open.

Kontinu: Jurnal Penelitian Didaktik Matematika
E-ISSN: 2656-5544
P-ISSN: 2715-7326
Vol. 10, No. 2: November 2026
Page 644-668

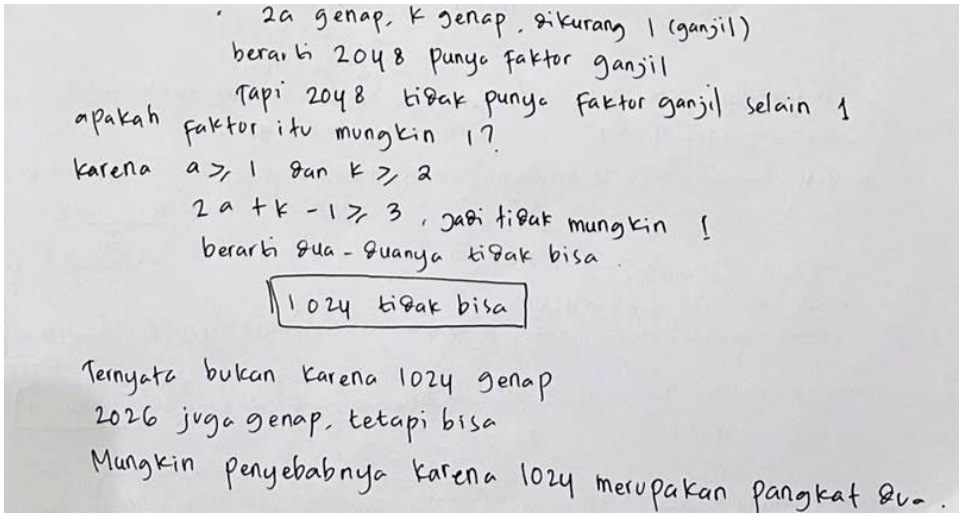


Figure 11. Excerpt from S3: the conclusion that 1024 is not representable and the tentative conjecture concerning powers of two.

Episode matrix and evidence-based coding

Table 3. Matrix of eleven observed episodes and their artifact-bounded coding interpretations. The tentative codes refer to written traces, not direct psychological assessments.

Episode	Observed written trace	Code	Mathematical appraisal and limits of inference
S1-E1	$x = 3, y = 3$ and then $x = 3$, $y = 6$ rejected; (3, 7, 42) obtained	PL, MN	Exploration yields a valid example; it does not yet establish all solutions.
S1-E2	$\frac{1}{2} \leq \frac{3}{x} \Rightarrow x \leq 6$; bounds on y and a formula for z derived	RG, EV	Explicitly narrows the cases; corrections and crossings-out are visible.
S1-E3	Noninteger z rejected; ten triples summarized	MN, EV	The final list is complete on the basis of the preceding bounds.
S2-E1	"$13 + 25 = 36$"; "But 38 is not 36"	MN	An inconsistency in the original relation is detected; internal thoughts and emotions are not observable.
S2-E2	Auxiliary line PE; $x = 2$; $y = \pm 3$; $y = -3$ rejected	PL, RG, EV	Representation is changed and the interior-point condition is used.
S2-E3	h = 5 or 1; BC = 5; PD = 2√2	EV (tentative)	Correct selection, although the reason for rejecting $h = 1$ is not fully written out.
S2-E4	"So perhaps" follows 13 + 20 = 25 + 8	PL/EV (tentative)	A general identity is conjectured; no proof for arbitrary rectangles is supplied.

Episode	Observed written trace	Code	Mathematical appraisal and limits of inference
S3-E1	Two- to four-term trials; a = 505; checking line 1011 + 1015 = 2020	PL, MN (limited)	The representation is correct, but numerical verification contains an arithmetic slip.
S3-E2	General k-term formula; divisors of 4052; $k \leq 63$	RG, EV	Completeness for part (a) is supported by bounds and divisibility.
S3-E3	$2048 = k(2a + k - 1)$; analysis of odd/even k	RG, EV	A valid impossibility proof for 1024 is provided.
S3-E4	"Perhaps the reason is ... a power of two"	PL (conjecturing)	The general conjecture in part (c) remains unproved.

Cross-case comparison and locally defined process profiles

Table 4. Comparison of obstacles, visible strategic responses, and the extent of mathematical achievement across three cases.

Case	Main obstacle	Visible strategic adjustment	Extent of achievement in the artifact
S1 (unit fractions)	Trying values cannot guarantee that all solutions have been found.	Bounds on x and y, a formula for z, and integer checks.	Ten triples and the basis for completeness are recorded.
S2 (geometry)	Direct use of the Pythagorean relationship is invalid/inconsistent.	Auxiliary line, two equations, and interpretation of signs and positions.	BC, PD, and P are correct; general identity checked numerically only.
S3 (consecutive sums)	The number of terms is initially unbounded; 1024 behaves differently from 2026.	k-term sum, factorization, and proof using odd/even cases.	Parts (a) and (b) are resolved; the power-of-two conjecture in (c) is unproved.

The three cases do not follow a uniform pathway. S1 moves from trying values to setting bounds that provide a criterion for completeness. S2 changes representations so that the selected theorem has a legitimate geometric basis. S3 uses the structure of products to replace an open-ended list of possible sequence lengths. A common feature is that progress becomes evident once work is guided by a constraint or a general property of the problem, rather than only by a successful numerical example. Because each student solved a different task, however, these contrasts cannot be used to rank students by metacognitive ability.

The written evidence also helps distinguish monitoring from evaluation. In S2, the note "38, not 36" marks detection of an inconsistency. The rejection of $y = -3$ evaluates whether a candidate solution satisfies the geometric constraints. In S1, the introduction of bounds on x and y signals a change in how the search is controlled, but its psychological motivation remains unknown. In S3, bounding k and reasoning about parity coexist with an error in a simple addition check. Regulation should therefore not be viewed as an ability displayed with equal strength throughout a student's work. A more defensible interpretation specifies what was effectively controlled—and what was not—in a given episode.

The comparison also guards against equating productivity with a flawless final solution. S1 achieves completeness, but the photographs do not reveal how long the difficulty lasted. S2 revises an invalid strategy and obtains correct dimensions, although the general proof is absent. S3 proves the impossibility for 1024 but ends with an unproved conjecture about all positive integers. Across these cases, the quality of the reasoning moves is more informative than a simple "success" or "failure" label. This reading is consistent with research showing that error detection and revision can be productive elements of mathematical activity (Granberg, 2016); (Warshauer, 2015); (Young et al., 2024).

Interpreting productive struggle and metacognitive regulation together

The two ideas are. Not the same, in analysis. A math problem and a good next step can show that a struggle was productive when looking at the written solution. Calling that step a thought-about-thinking move needs a guess: a crossed-out calculation might show that something does not match, while a new way of writing might show a plan is being changed but neither shows that the student truly knows. This difference stops the study from saying all mistakes are struggle or that all fixes prove thought-about-thinking. It also lets the three cases be compared without assuming the tasks test the mental skills.

Answers to the research questions and analytical contribution

Three visible challenges deal with the research questions. S1 deals with the need to create a full solution set; S2 deals, with replacing a relationship that does not have a right-

angle condition; and S3 deals with telling unsuccessful trials of short sequences from an argument that covers every possible sequence length. Written actions that could be connected to regulation come in different forms: setting bounds on variables adding an extra line rejecting values that are not allowed breaking numbers into factors and checking if numbers are even or odd. The strongest idea that applies across cases is not that the students have fixed psychological "profiles " but that each task has moments where the students' written methods become more limited or more mathematically supported.

The studys main contribution is to separate four types of evidence that are usually mixed up: finding an error or something that doesn't match fixing it making progress in math and giving proof. These can happen in amounts. S2 has a correction in representation but no overall proof; S3 has a strong proof, for one case but doesn't finish the general idea. This episode-based reading is more consistent with Schoenfeld's (1985) mathematical control and Flavell's (1979) cognitive regulation than assigning a broad label of "good metacognition" on the basis of one answer (SCHOENFELD, 1985); (Flavell, 1979). Such evidence can instead inform interviews designed to investigate the actual reasons behind strategic changes.

Limitations and implications

Three limitations are particularly important. First, verified information is unavailable about the circumstances of task completion, prior instruction, time allocation, or teacher intervention. Second, photographs preserve written traces but not the complete sequence of actions, spontaneous speech, or affective experiences. Third, each student is represented by only one task, making it impossible to separate student characteristics from task characteristics. Accordingly PL, MN, EV and RG are categories that are supported by the written products. PL, MN, EV and RG are not diagnoses of competence. If you want to publish this work as field research you should document the participation procedures permission to use the artifacts and the coding audit. Triangulation can only be claimed when interview or recording data are available.

For teaching, these three tasks suggest questions that preserve cognitive demand: "What justifies stopping the search?", "Is the triangle you are using actually right-angled?", and "Does one example prove the claim for every case?" Such prompts do not

provide the solution; instead, they invite students to check the limits and strength of their own reasoning. For further research, the photographs offer useful interview stimuli: why S1 began to bound x, why S2 abandoned the original equation, and what led S3 from evenness to powers of two. Those answers must be elicited from the students, not supplied through invented dialogue.

CONCLUSION

Analysis of six photographs of three students' work reveals three forms of progress in response to nonroutine mathematical tasks. S1 bounds the search and records all ten valid triples. S2 corrects an invalid application of Pythagoras, builds an appropriate representation, and obtains BC = 5, PD = $2\sqrt{2}$, and P = (2, 3); the identity $PA^2 + PC^2 = PB^2 + PD^2$ remains a conjecture based on a specific example in the student's work. S3 reveals the way to write 2026 with consecutive numbers shows that 1024 cannot be split into two or more consecutive positive integers and suggests—without proof—a broad rule about powers of two. In each case the idea of struggle can be best guessed from how the written solution changes after a math limit appears. Still photographs cannot show what students think while they work. These findings are just interpretations of artifacts. They need support, from recorded fieldwork, interviews and real recordings before we can say they apply widely.

REFERENCES

Amalina, I. K., & Vidákovich, T. (2023). Development and differences in mathematical problem-solving skills: A cross-sectional study of differences in demographic backgrounds. *Heliyon*, *9*(5), e16366. https://doi.org/10.1016/j.heliyon.2023.e16366

Aydın, U., & Özgeldi, M. (2024). What's metacognition got to do with the relationship between test anxiety and mathematics achievement? *European Journal of Psychology of Education*, *39*(3), 2509–2529. https://doi.org/10.1007/s10212-024-00797-7

Baumanns, L., & Rott, B. (2023). Identifying Metacognitive Behavior in Problem-Posing Processes. *International Journal of Science and Mathematics Education*, *21*(5), 1381–1406. https://doi.org/10.1007/s10763-022-10297-z

Biccard, P. (2024). Productive Struggle in Mathematical Modelling. *The Mathematics Enthusiast*, *21*(1), 99–112. https://doi.org/10.54870/1551-3440.1620

Braithwaite, D. W., & Sprague, L. (2021). Conceptual Knowledge, Procedural Knowledge, and Metacognition in Routine and Nonroutine Problem Solving. *Cognitive Science*, *45*(10), e13048. https://doi.org/10.1111/cogs.13048

Del Olmo-Muñoz, J., Diago, P. D., Arnau, D., Arnau-Blasco, D., & González-Calero, J. A. (2024). Metacognitive control during problem solving at early ages in programming tasks using a floor robot. *ZDM – Mathematics Education*, *56*(6), 1303–1317. https://doi.org/10.1007/s11858-024-01621-8

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data, Rev. ed* (p. liii, 443). The MIT Press.

Favier, S., & Dorier, J.-L. (2024). Heuristics and semantic spaces for the analysis of students' work in mathematical problem solving. *Educational Studies in Mathematics*, *115*(3), 407–431. https://doi.org/10.1007/s10649-023-10297-y

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Granberg, C. (2016). Discovering and addressing errors during mathematics problem-solving—A productive struggle? *The Journal of Mathematical Behavior*, *42*, 33–48. https://doi.org/10.1016/j.jmathb.2016.02.002

Ihsani, U. A., & Siswono, T. Y. E. (2024). Student's Metacognition in Solving Mathematical Problem using Chat-GPT. *Journal of Mathematical Pedagogy (JoMP)*, *5*(2), 102–112. https://doi.org/10.26740/jomp.v5n2.p102-112

Jin, S., & Leong, Y. H. (2025). Metacognition in mathematics problem solving: A case of making monitoring and regulation visible. *The Mathematician Educator*. https://hdl.handle.net/10497/30992

Khalil, I., AL-Otaibi, A., Almughyirah, S., & Almalky, M. (2025). How do students evaluate their teachers' support for productive struggle in learning mathematics? A mixed-methods study. *Cogent Education*, *12*(1), 2442234. https://doi.org/10.1080/2331186X.2024.2442234

Kholid, M. N., & Ahadiyati, A. (2022). Students' metacognition in solving non-routine problems. *Al-Jabar : Jurnal Pendidikan Matematika*, *13*(1), 125–138. https://doi.org/10.24042/ajpm.v13i1.11776

Kusaka, S., & Claude Habimana, J. (2025). Metacognitive skills in collaborative problem solving: Case of secondary mathematics in Rwanda. *The Journal of Educational Research*, *118*(4), 368–375. https://doi.org/10.1080/00220671.2025.2486473

Kusaka, S., & Ndihokubwayo, K. (2022). Metacognitive strategies in solving mathematical word problems: A case of Rwandan primary school learners. *SN Social Sciences*, *2*(9), 186. https://doi.org/10.1007/s43545-022-00495-5

Mufliva, R., Turmudi, T., Herman, T., Iriawan, S. B., & Fitriani, A. D. (2024). Problem-Solving Skills and Productive Struggle of Students in Solving Mathematical Problems in Elementary School. *Southeast Asian Mathematics Education Journal*, *14*(1), 69–82. https://doi.org/10.46517/seamej.v14i1.376

Muncer, G., Higham, P. A., Gosling, C. J., Cortese, S., Wood-Downie, H., & Hadwin, J. A. (2022). A Meta-Analysis Investigating the Association Between Metacognition and Math Performance in Adolescence. *Educational Psychology Review*, *34*(1), 301–334. https://doi.org/10.1007/s10648-021-09620-x

Oudman, S., van de Pol, J., & van Gog, T. (2022). Effects of self-scoring their math problem solutions on primary school students' monitoring and regulation. *Metacognition and Learning*, *17*(1), 213–239. https://doi.org/10.1007/s11409-021-09281-9

Prastiti, T. D., Ahmadah, I. F., & Siswono, T. Y. E. (2026). Metacognitive regulation in collaborative math problem-solving among heterogeneous secondary students. *Eurasia Journal of Mathematics, Science and Technology Education*, *22*(5), em2834. https://doi.org/10.29333/ejmste/18518

Rajadurai, R., & Ganapathy, H. (2023). Effect of use of metacognitive instructional strategies in promoting mathematical problem solving competence amongst undergraduate students in facing competitive examination. *Cogent Social Sciences*, *9*(1), 2173103. https://doi.org/10.1080/23311886.2023.2173103

Ramlah, R., Siswono, T. Y. E., & Lukito, A. (2024). Revealing the uniqueness of variations in prospective teachers' metacognitive activities in solving mathematical problems based on gender. *Infinity Journal*, *13*(2), 477–500. https://doi.org/10.22460/infinity.v13i2.p477-500

Scheibe, D. A., Was, C. A., Dunlosky, J., & Thompson, C. A. (2023). Metacognitive Cues, Working Memory, and Math Anxiety: The Regulated Attention in Mathematical Problem Solving (RAMPS) Framework. *Journal of Intelligence*, *11*(6), 117. https://doi.org/10.3390/jintelligence11060117

SCHOENFELD, A. H. (1985). *Mathematical Problem Solving*. Elsevier Science.

Schonberg, C., Hargis Becker, M., An, X., & Bashkov, B. M. (2026). Spaced Learning Supports Productive Struggle in an Online Learning Platform. *Technology, Knowledge and Learning*, *31*(1), 365–381. https://doi.org/10.1007/s10758-025-09914-x

Torregrosa, A., & Albarracín, L. (2026). Characterizing metacognitive processes and mathematical skills during problem solving by using a non-linear orientation base as a self-regulation tool. *Educational Studies in Mathematics*, *121*(2), 333–364. https://doi.org/10.1007/s10649-025-10451-8

Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, *18*(4), 375–400. https://doi.org/10.1007/s10857-014-9286-3

Young, J., Bevan, D., & Sanders, M. (2024). How Productive is the Productive Struggle? Lessons Learned from a Scoping Review. *International Journal of Education in Mathematics, Science and Technology*, *12*(2), 470–495. https://doi.org/10.46328/ijemst.3364