ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 7 (July 2026)

Mathematical Reasoning Gaps and Problem-Solving Confidence among Grade 6 Learners

Filamer V. Tarun[1*] and Michelle S. Morales[1,2]
[1]*Northeastern College*
[*]*filamertarun45@gmail.com*, [2]*michellemorales@northeasterncollege.edu.ph*

Date Submitted:
February 5, 2026

Date Accepted:
April 13, 2026

Date Published:
July 17, 2026

DOI:
10.5281/zenodo.21403749

ABSTRACT

This study traced how mathematical reasoning gaps were associated with problem-solving confidence among Grade 6 learners in a public elementary school in the City of Ilagan, Isabela. A cross-sectional diagnostic relational design was used. Data were collected through the Mathematical Reasoning Gap Assessment and the Problem-Solving Confidence Inventory, both of which underwent expert validation, pilot testing, and reliability analysis. Partial credit Rasch modeling was applied to identify reasoning difficulties, while a Rasch rating scale model measured confidence. Robust regression, Spearman correlation, and a reasoning-confidence calibration index were also employed. Results showed an overall moderate mathematical reasoning gap, with the greatest weaknesses found in logical justification and answer verification, followed by mathematical representation and strategy selection. Problem-solving confidence was at a developing level, with the lowest confidence observed in explaining solutions and persisting after an unsuccessful attempt. Confidence significantly predicted mathematical reasoning ability, although it explained only a modest portion of the variance. Calibration findings further revealed that several learners either overestimated or underestimated their actual reasoning performance. The findings indicated that mathematical difficulty was not confined to computation but also involved interpretation, strategy use, explanation, verification, and self-judgment. Mathematics instruction should therefore include diagnostic reasoning tasks, structured justification, multiple representations, explicit verification strategies, opportunities for solution revision, and feedback that helps learners align confidence with demonstrated competence.

Keywords: *answer verification, Grade 6 learners, mathematical reasoning, problem-solving confidence, Rasch analysis, strategy selection*

INTRODUCTION

Mathematics becomes most meaningful when learners can explain why a solution works, examine whether an answer is reasonable, and choose an appropriate strategy when a familiar procedure no longer fits. These abilities go beyond accurate computation. They involve recognizing patterns, forming conjectures, making comparisons, generalizing relationships, and supporting conclusions with valid mathematical ideas. Jeannotte and Kieran described mathematical reasoning as having both structural and process dimensions, which means that reasoning can be seen in the way learners organize mathematical ideas and in the thinking processes they use to reach and justify conclusions. When these dimensions are not sufficiently developed, learners may complete routine exercises but struggle when a problem requires interpretation, explanation, or a different solution path (Jeannotte & Kieran, 2017).

Evidence from primary education in the Philippines shows that many learners reach the later elementary grades without attaining higher levels of mathematical proficiency. The 2024 Southeast Asia Primary Learning

ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 7 (July 2026)

Metrics assessed 5,070 Grade 5 learners from the country and found that the number reaching Band 6 or higher increased from 17 out of 100 learners in 2019 to 26 out of 100 in 2024. Although this improvement is encouraging, most Grade 5 learners remained below the higher proficiency bands. The assessment also reported continuing differences between high and low performers and disparities associated with sex, household resources, language used at home, and geographic location. Since Grade 6 immediately follows this stage, uneven mastery acquired in the earlier grades may appear as difficulty in interpreting problems, connecting concepts, and explaining mathematical decisions during more demanding lessons (United Nations Children's Fund [UNICEF], 2025).

The persistence of mathematical difficulty is also visible in assessments conducted at the secondary level. In the Programme for International Student Assessment 2022, only 16 percent of Filipino students reached at least Level 2 proficiency in mathematics, compared with the Organisation for Economic Co-operation and Development average of 69 percent. Learners at Level 2 can recognize how a simple real-life situation may be represented mathematically even when direct instructions are not given. Almost no Filipino participants reached the highest proficiency levels, where students are expected to model complex situations and evaluate appropriate problem-solving strategies. These findings do not identify the specific reasoning gaps of Grade 6 learners, but they suggest that weaknesses in applying, interpreting, and evaluating mathematical ideas may continue when they are not addressed during the elementary years (Organisation for Economic Co-operation and Development [OECD], 2023).

One possible source of this difficulty is an excessive dependence on remembered procedures. Learners may know which operation is usually associated with a particular type of exercise, yet remain uncertain when the wording, numerical information, or structure of the problem changes. This condition can produce correct answers in repetitive tasks without establishing a strong understanding of the relationships behind the procedure. Jonsson et al. (2020) found that students who practised creative mathematical reasoning performed better on both practised and transfer tasks than those who relied mainly on algorithmic reasoning. Their findings indicate that opportunities to construct and justify solutions can support understanding that extends beyond the original exercise. For Grade 6 learners, such opportunities are important because the curriculum increasingly expects them to work with multistep problems, fractions, decimals, percentages, ratios, measurement, geometry, and data in situations where the required operation may not be immediately evident.

Difficulty in mathematical reasoning also has an affective side. A learner who repeatedly fails to understand a problem may begin to expect failure before attempting the next one. Problem-solving confidence refers to the learner's belief that he or she can understand a mathematical task, select a workable approach, persist through difficulty, and arrive at a defensible answer. It does not simply mean feeling positive about mathematics. Confidence is connected with the learner's willingness to engage with the task and continue working when an initial strategy is unsuccessful. A review of mathematics self-efficacy research found that students' beliefs about their mathematical capability were strongly associated with learning behavior, motivation, effort, perseverance, and performance. The review also emphasized that mathematical self-efficacy is dynamic and may differ according to the topic, task, level of difficulty, classroom, and school setting (Street et al., 2024).

The relationship between perceived difficulty and confidence is especially important for learners in Grade 6. Nuutila et al. (2021) studied 1,024 sixth-grade learners during a dynamic problem-solving activity and observed that self-efficacy and interest tended to decline as the task continued, while perceived difficulty initially increased. Learners' starting levels of interest and their changes in self-efficacy were associated with task performance. The findings show that confidence is not fixed throughout problem solving. It may weaken while a learner confronts unfamiliar information, failed attempts, or increasing task demands. A Grade 6 learner may therefore possess some relevant mathematical knowledge but still withdraw from a problem because the task is perceived as too difficult or because earlier attempts have reduced confidence.

Confidence may also vary according to the particular content being taught and the level of challenge presented. In a study involving Norwegian Grade 6 and Grade 10 students, learners reported their self-efficacy while studying new mathematical topics and completing tasks with easy, moderate, and difficult levels. The study showed that self-efficacy could change within a short sequence of lessons rather than remaining stable throughout

ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 7 (July 2026)

a school year. This finding is relevant to Grade 6 classrooms because a learner may appear confident when completing basic calculations but become uncertain when asked to apply the same concept in a nonroutine word problem. Measuring only general confidence may therefore conceal the specific situations in which learners hesitate, avoid reasoning, or depend on teacher guidance (Street et al., 2022).

Large-scale evidence further supports the need to examine mathematical capability and confidence together. An analysis of the 2019 National Assessment of Educational Progress found that mathematics self-efficacy was a significant predictor of achievement in Grades 4, 8, and 12. The relationship was present even after student and school characteristics were considered, and several achievement gaps became smaller when differences in self-efficacy were taken into account. These findings suggest that low achievement should not be viewed only as a lack of knowledge. Learners' judgments about their own capability may influence whether they attempt difficult problems, sustain effort, revise an incorrect strategy, and use the reasoning skills they possess (Yang et al., 2024).

Grade 6 represents a significant point for examining these concerns because learners are completing elementary education and preparing for the greater abstraction and independence required in secondary mathematics. Unidentified reasoning gaps at this level may affect later learning, while weak problem-solving confidence may discourage learners from attempting tasks that could strengthen their understanding. The study determine the areas of mathematical reasoning in which learners experience difficulty, describe their confidence when solving mathematical problems, and examine whether these two conditions are significantly related. The findings may provide a sound basis for instructional support that responds not only to incorrect answers but also to the reasoning processes, uncertainties, and problem-solving behaviors behind them.

Literature Review

Identification of Mathematical Reasoning Gaps

Mathematical reasoning gaps are not limited to incorrect answers because learners may encounter difficulty at different points in the solution process. Some learners fail to understand the mathematical conditions presented in a task, while others cannot recall relevant concepts, select a suitable strategy, carry out the necessary operations, or justify the conclusion they have reached. Säfström et al. (2024) developed a diagnostic framework from 285 examples of reasoning difficulties involving learners from Grades 1 to 12. The framework emphasized the importance of identifying the specific point at which a learner's reasoning becomes unsupported rather than treating all unsuccessful solutions as evidence of the same weakness. Such an approach is important for Grade 6 learners because a single problem may require several connected actions, including interpreting information, representing relationships, making calculations, and explaining why the answer is valid. A learner may succeed in one action but struggle in another. Careful diagnosis therefore allows teachers to distinguish conceptual misunderstanding from procedural error, incomplete reasoning, and difficulty in communicating mathematical ideas. This distinction can support instructional responses that address the actual source of difficulty without removing the learner's responsibility for developing the solution.

Linguistic and Numerical Demands of Mathematical Problems

The wording and structure of a mathematical problem can influence whether elementary learners recognize the relationships needed for a solution. Word problems require learners to coordinate reading comprehension, numerical knowledge, working memory, and reasoning, which means that weak performance cannot always be attributed to computation alone. Vessonen et al. (2024) reviewed 69 studies involving elementary school learners from Grades 1 to 6 and found that problem-solving performance was associated with several linguistic and numerical characteristics. Learners experienced greater difficulty when problems contained irrelevant information, inconsistent wording, unfamiliar placement of the unknown quantity, realistic conditions that had to be considered, or several mathematical operations. These findings show that a Grade 6 learner may know how to perform an operation but still fail to determine when and how it should be applied. The result may appear to be a gap in mathematical knowledge even when the difficulty begins with interpreting the problem statement or organizing its information. Assessing mathematical reasoning should therefore include tasks with varied structures so that teachers can determine whether learners can move beyond familiar keywords and identify the mathematical relationships represented in different forms.

Metacognitive Regulation in Mathematical Problem Solving

Successful problem solving requires learners to manage their thinking while they work rather than apply a strategy without checking its suitability. Metacognitive regulation includes planning how to approach a task, monitoring whether the chosen steps are producing meaningful results, and evaluating the reasonableness of the final answer. Xie et al. (2024) conducted a meta-analysis of 147 studies with 338 independent samples and found a significant positive relationship between metacognition and mathematics achievement. The relationship varied according to age, mathematical domain, and cultural setting, while mathematical forms of metacognition were more closely connected with mathematics achievement than general metacognitive ability. For Grade 6 learners, limited monitoring may be seen when they continue an ineffective procedure, overlook contradictory information, accept an unreasonable answer, or cannot explain why they selected a certain operation. These behaviors reveal a reasoning gap that may remain hidden when assessment focuses only on the final numerical response. Learners need opportunities to state their plans, compare strategies, inspect intermediate results, and revise an approach when evidence shows that it is not working. Examining these actions can provide a clearer account of how learners regulate mathematical reasoning during problem solving.

Problem-Solving Confidence and Emotional Responses

Problem-solving confidence reflects a learner's belief that a mathematical task can be understood and completed through appropriate effort and strategy. It affects the willingness to begin a difficult problem, remain engaged after an error, seek another method, and explain a solution. Zakariya's (2022) systematic review showed that mathematics self-efficacy can be strengthened through interventions that address mastery experiences, constructive feedback, peer modeling, emotional responses, teaching methods, and learning strategies. Confidence is therefore shaped by classroom experiences rather than treated as a fixed characteristic of the learner. Di Leo and Muis (2020) also found that Grade 5 learners who received cognitive-emotional strategy training performed better in complex mathematical problem solving, used more effective cognitive and metacognitive strategies, and managed confusion more successfully than learners who did not receive the intervention. These findings are relevant to Grade 6 because confusion and failed attempts may either encourage further reasoning or lead to frustration and withdrawal. A learner with low confidence may avoid unfamiliar problems even when the required knowledge is partly available, while a confident learner may persist long enough to identify and correct a reasoning error. Mathematical reasoning gaps and problem-solving confidence should therefore be examined together because cognitive difficulty and emotional response may influence each other during task engagement.

METHODS

Research Design

The study employed a cross-sectional diagnostic relational design. This design combined a detailed identification of mathematical reasoning gaps with an examination of how these gaps corresponded with the learners' confidence in solving mathematical problems. The diagnostic component determined the specific reasoning processes in which learners experienced difficulty, including problem interpretation, mathematical representation, strategy selection, logical justification, and answer verification. The relational component examined whether differences in problem-solving confidence were associated with variations in demonstrated mathematical reasoning. The design was appropriate because the investigation did not introduce an intervention or manipulate classroom conditions. Instead, it measured the learners' existing reasoning performance and confidence during the same period and analyzed the alignment between these two constructs. This design also allowed the study to distinguish learners who demonstrated both low reasoning and low confidence from those whose confidence did not accurately correspond with their actual reasoning performance.

Research Locale

The investigation was conducted at Sta. Isabel Sur Elementary School, a public elementary school under the Schools Division Office of the City of Ilagan, Isabela. The school served learners from Sta. Isabel Sur and nearby communities and implemented the prescribed Grade 6 mathematics curriculum of the Department of Education. The locale was selected because mathematical problem solving formed a regular part of Grade 6 instruction, while differences in learners' ability to explain procedures, identify relationships, and sustain effort during challenging tasks had been observed in classroom activities. The school provided an appropriate setting for examining reasoning gaps and problem-solving confidence under regular instructional conditions. Data collection was scheduled without disrupting major examinations, school programs, or required classroom activities.

Participants and Sampling Technique

The participants were officially enrolled Grade 6 learners at Sta. Isabel Sur Elementary School during the period of data collection. Learners were included when they had received parental or guardian consent, had provided their own assent, and were present during the scheduled administration of the research instruments. Learners who did not complete either of the two instruments were excluded from the paired analysis because their reasoning and confidence scores could not be properly matched.

Proportionate stratified random sampling was applied, with each Grade 6 class section treated as a sampling stratum. The number selected from each section was determined according to its proportion of the total Grade 6 enrollment. Within each stratum, participants were selected through computer-generated random numbers based on the official class lists. This procedure ensured that every Grade 6 section was represented without allowing one class to contribute a disproportionately large share of the participants. Replacement selection was conducted only when a selected learner did not meet the consent or attendance requirements.

Research Instrument

Data were gathered through a researcher-developed assessment package composed of the Mathematical Reasoning Gap Assessment and the Problem-Solving Confidence Inventory. The Mathematical Reasoning Gap Assessment contained selected-response and constructed-response tasks based on the Grade 6 mathematics competencies prescribed by the Department of Education. The tasks covered number relationships and operations, fractions and decimals, ratio and percentage, geometry and measurement, patterns, and data interpretation. Instead of measuring computation alone, the assessment required learners to interpret given information, represent mathematical relationships, choose and carry out a strategy, justify a conclusion, and examine whether the answer was reasonable. Constructed responses were scored using an analytic rubric that differentiated complete reasoning, partially supported reasoning, procedural work without justification, and unsupported or incorrect responses.

The Problem-Solving Confidence Inventory measured confidence in beginning an unfamiliar problem, interpreting mathematical information, choosing a strategy, continuing after an unsuccessful attempt, explaining a solution, checking an answer, and solving tasks without immediate teacher assistance. The statements were answered through a four-point response scale ranging from not confident to highly confident. A neutral category was not included so that the learners indicated the direction that most closely represented their actual confidence. The wording remained brief and suitable for Grade 6 reading ability.

The initial instruments were examined by a validation panel composed of mathematics education specialists, an educational measurement expert, and an experienced Grade 6 mathematics teacher. The validators assessed the relevance, clarity, developmental appropriateness, curricular alignment, and adequacy of each item. Items that contained unnecessary language, unclear mathematical conditions, or more than one possible interpretation were revised or removed. The item-level content validity indices ranged from .83 to 1.00, while the scale-level content validity index based on the average method was .94. These values indicated strong agreement among the validators regarding the suitability of the instruments.

ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 7 (July 2026)

The revised instruments were pilot-tested with Grade 6 learners from a comparable public elementary school that was not included in the main investigation. The pilot administration examined completion time, clarity of directions, response patterns, item difficulty, discrimination, and the functioning of the confidence response categories. The Mathematical Reasoning Gap Assessment obtained a provisional Cronbach's alpha coefficient of .86. The Problem-Solving Confidence Inventory produced a provisional overall Cronbach's alpha coefficient of .92, while its domain coefficients ranged from .83 to .89. Double scoring of the constructed-response items produced an intraclass correlation coefficient of .91, indicating a high level of consistency between raters. Items with weak discrimination, unclear scoring boundaries, or redundant content were revised before the final administration.

Data Gathering

Approval to conduct the study was secured from the appropriate school and education authorities before any learner was approached. The researcher coordinated with the school head and Grade 6 teachers regarding suitable dates, available rooms, and the length of each administration. Parent or guardian consent forms and learner assent forms were distributed before data collection. The forms described the purpose of the study, the activities required, the voluntary nature of participation, and the measures used to protect confidentiality.

The Mathematical Reasoning Gap Assessment was administered first under standardized classroom conditions. The same directions, time allowance, and clarification rules were applied across all participating sections. Learners completed the assessment independently, and teachers did not provide hints or confirm whether an answer was correct. A short break was provided before the administration of the Problem-Solving Confidence Inventory. The confidence inventory was completed after the reasoning assessment so that learners based their judgments on an immediate problem-solving experience rather than on a vague impression of their mathematics ability.

Each participant received a coded identifier that was used on both instruments. Names were not written on the answer sheets used for analysis. Completed instruments were checked only for missing pages and unanswered sections, without altering any learner response. Constructed responses were scored independently by two trained raters using the approved analytic rubric. Differences between the raters were reviewed through the scoring descriptors, and a final score was assigned only after agreement had been reached. The encoded data were examined against the original answer sheets to reduce data-entry errors.

Data Analysis

The analysis began with data screening to identify incomplete paired records, encoding errors, extreme response patterns, and violations that could affect the selected statistical procedures. Frequencies, medians, interquartile ranges, means, and standard deviations were used to describe item and domain performance. The percentage of learners who demonstrated incomplete or unsupported reasoning was also calculated for each assessed reasoning process.

A partial credit Rasch model was applied to the Mathematical Reasoning Gap Assessment. This model was selected because the constructed-response tasks contained ordered score categories rather than simple correct or incorrect scoring. The analysis generated item difficulty estimates and learner ability estimates expressed in logits. Item fit was examined through infit and outfit mean-square statistics, with values from 0.70 to 1.30 treated as productive for measurement. A Wright person-item map was used to compare the distribution of learner ability with the difficulty of the reasoning tasks. Mathematical reasoning gaps were identified when the difficulty of an item or reasoning domain exceeded the average ability estimate of the learners and when a substantial proportion of responses showed incomplete reasoning.

A Rasch rating scale model was applied to the Problem-Solving Confidence Inventory. The analysis examined whether the response categories progressed in the intended order and whether each statement contributed meaningfully to the measurement of confidence. Learner confidence estimates were converted into interval-level logit measures rather than being interpreted solely through raw Likert averages.

ISSN: 3116-3475
International Journal
of Education, Research, and Innovation Perspectives
Volume 2 Issue 7 (July 2026)

A reasoning-confidence calibration index was computed by standardizing each learner's Rasch reasoning ability measure and confidence measure. The standardized reasoning score was subtracted from the standardized confidence score. Values close to zero indicated that confidence was generally aligned with demonstrated reasoning. Positive values indicated confidence that exceeded demonstrated reasoning, while negative values indicated confidence that was lower than the learner's measured reasoning performance. This analysis allowed the study to identify confidence miscalibration that could not be detected through a conventional correlation alone.

The association between mathematical reasoning and problem-solving confidence was examined through robust regression using Huber M-estimation. This procedure was selected because it reduced the influence of unusually high or low scores and remained dependable when the residuals did not fully satisfy ordinary least squares assumptions. Problem-solving confidence served as the explanatory variable, while mathematical reasoning ability served as the outcome variable. Bootstrap confidence intervals based on repeated resampling were generated for the regression coefficient. Spearman's rank-order correlation was also computed as a sensitivity analysis to determine whether the direction and strength of the association remained consistent without assuming normally distributed scores. Statistical significance was evaluated at the .05 level, while effect sizes and confidence intervals were given greater attention than the probability value alone.

Ethical Consideration

The study followed the ethical principles of voluntary participation, informed permission, learner protection, confidentiality, and responsible data management. Written consent was obtained from parents or guardians, while assent was obtained from the participating learners using language appropriate for their age. The learners were informed that participation was optional and that they could stop at any point without receiving a penalty. Their decision to participate or decline did not affect their mathematics grades, class standing, access to school activities, or relationship with their teachers.

The assessment was not presented as a graded examination. Learners were not publicly identified according to their reasoning scores or confidence levels. Individual results were not used to label learners as weak, incapable, overconfident, or underconfident. Only summarized findings were reported in the manuscript. Electronic data were stored in password-protected files, while printed materials were kept in a secured location accessible only to the researcher and authorized research personnel. Coded data were retained only for the period required for analysis, verification, and publication and were disposed of according to the approved data management procedure. Findings were reported accurately, including results that did not support the expected relationship between mathematical reasoning gaps and problem-solving confidence.

RESULTS AND DISCUSSION

Table 1. *Rasch Measurement Statistics for Mathematical Reasoning and Problem-Solving Confidence*

Measurement Indicator	Mathematical Reasoning Gap Assessment	Problem-Solving Confidence Inventory
Mean person measure, logits	-0.47	-0.18
Standard deviation of person measures	0.86	0.79
Person reliability	0.84	0.88
Person separation index	2.31	2.68
Item reliability	0.96	0.94
Item separation index	4.87	3.96
Item infit mean-square range	0.78 to 1.22	0.81 to 1.19
Item outfit mean-square range	0.74 to 1.27	0.77 to 1.25
Cronbach's alpha	0.88	0.93
Ordered response thresholds	Not applicable	Yes

Measurement Indicator	Mathematical Reasoning Gap Assessment	Problem-Solving Confidence Inventory
Overall model interpretation	Productive measurement	Productive measurement

The Rasch statistics indicated that both instruments functioned adequately during the main administration. The person reliability coefficient of 0.84 for mathematical reasoning showed that the assessment consistently differentiated learners according to their reasoning ability. Its person separation index of 2.31 suggested that the participants could be classified into approximately three distinguishable levels of reasoning performance. The item reliability of 0.96 and separation index of 4.87 showed that the ordering of task difficulty was stable and was unlikely to have resulted from chance variation.

The Problem-Solving Confidence Inventory produced a person reliability coefficient of 0.88 and a separation index of 2.68. These findings indicated that the inventory could distinguish learners with low, developing, and relatively strong confidence. The item reliability coefficient of 0.94 confirmed that the confidence statements had sufficiently different endorsement difficulty levels. All infit and outfit statistics remained within the prespecified acceptable interval of 0.70 to 1.30. No task or statement showed severe underfit or overfit, which meant that the items contributed productively to their intended measures.

The main-administration Cronbach's alpha coefficients of 0.88 for mathematical reasoning and 0.93 for problem-solving confidence were slightly higher than those recorded during the pilot test. The improvement suggested that the revisions made after pilot testing had strengthened item clarity and internal consistency. The ordered thresholds of the four confidence categories also showed that learners distinguished among not confident, slightly confident, confident, and highly confident rather than using the categories randomly.

Table 2. *Rasch Difficulty Estimates and Mathematical Reasoning Gaps by Domain*

Mathematical Reasoning Domain	Domain Difficulty, Logits	Gap Index	Incomplete or Unsupported Responses, %	Gap Classification	Rank
Logical justification	0.76	1.23	71.20	Severe gap	1
Answer verification	0.49	0.96	64.90	Severe gap	2
Mathematical representation	0.21	0.68	57.80	Moderate gap	3
Strategy selection	0.08	0.55	53.40	Moderate gap	4
Problem interpretation	-0.18	0.29	44.60	Emerging gap	5
Overall mathematical reasoning	0.00	0.47	58.40	Moderate gap	

Note. The gap index was calculated by subtracting the mean learner ability estimate of -0.47 logits from the domain difficulty estimate. A positive index indicated that the reasoning demand exceeded the average measured ability. Gap classifications were defined as emerging from 0.00 to 0.49, moderate from 0.50 to 0.89, and severe at 0.90 or higher.

The learners obtained a mean mathematical reasoning ability of -0.47 logits, which was below the Rasch item mean of zero. This difference showed that the assessment tasks were generally more demanding than the reasoning level demonstrated by the average learner. The overall gap was classified as moderate, with 58.40 percent of the responses containing incomplete, procedural, or unsupported reasoning. Learners frequently obtained partial credit because they identified relevant numerical information or performed an operation but did not complete the reasoning needed to establish the validity of their answers.

Logical justification emerged as the most serious area of difficulty. Its domain difficulty of 0.76 logits was 1.23 logits above average learner ability, and 71.20 percent of the responses lacked a complete explanation. A common response pattern consisted of writing an equation and numerical answer without explaining why the selected operation represented the relationship described in the problem. Some learners repeated parts of the problem statement as their explanation, while others offered statements such as "because I divided" without establishing why division was mathematically appropriate.

Answer verification ranked second, with a gap index of 0.96 logits. Almost two thirds of the responses did not show a meaningful checking process. Learners commonly recomputed the same operation in the same way, copied the original answer, or wrote "checked" without presenting evidence. Only a smaller group used estimation, inverse operations, substitution, comparison with known quantities, or examination of the answer's reasonableness. This result indicated that obtaining an answer was often treated as the end of the task rather than as a claim that required evaluation.

Mathematical representation and strategy selection showed moderate gaps. Learners encountered difficulty in converting verbal information into diagrams, tables, number sentences, models, or organized lists. Some selected an operation from a familiar keyword without examining the complete relationship among the quantities. Such responses appeared more often in multistep problems and tasks involving fractions, percentages, measurement, and comparison. Pongsakdi et al. found that difficult word problems demand both arithmetic competence and deeper comprehension of the situation represented by the text. Their findings help explain why performing individual operations correctly did not guarantee success when the Grade 6 learners had to form an accurate representation and coordinate several solution steps (Pongsakdi et al., 2020).

Problem interpretation had the smallest gap, but 44.60 percent of responses still showed incomplete identification of the given information, required outcome, or mathematical condition. The result suggested that many learners could extract visible numbers but had difficulty deciding which information was relevant and how the quantities were related. The difficulty was therefore not confined to computation. It began during the construction of meaning from the problem and continued through representation, strategy use, justification, and verification.

Table 3. *Rasch Measures of Problem-Solving Confidence by Domain*

Problem-Solving Confidence Domain	Mean Measure, Logits	Standard Deviation	Interpretation	Rank
Checking whether an answer is correct	0.14	0.82	Developing confidence	1
Interpreting mathematical information	-0.09	0.77	Developing confidence	2
Solving without immediate teacher assistance	-0.11	0.83	Developing confidence	3
Choosing an appropriate strategy	-0.24	0.80	Developing confidence	4
Continuing after an unsuccessful attempt	-0.36	0.91	Low confidence	5
Explaining and defending a solution	-0.41	0.88	Low confidence	6
Overall problem-solving confidence	-0.18	0.79	Developing confidence	

Note. Measures below -0.30 logits were interpreted as low confidence, measures from -0.29 to 0.29 as developing confidence, and measures of 0.30 or higher as strong confidence. The Rasch category thresholds were ordered at -1.24, -0.16, and 1.40 logits. Category outfit mean-square values ranged from 0.93 to 1.08.

The overall confidence measure of -0.18 logits indicated that the learners had developing rather than strong problem-solving confidence. They generally believed that they could attempt common Grade 6 mathematics tasks, but their confidence weakened when a problem required independent decision-making, explanation, or recovery from an unsuccessful strategy. None of the measured domains reached the strong-confidence range.

Confidence in checking answers received the highest estimate at 0.14 logits. This result initially appeared inconsistent with the severe reasoning gap found in actual answer verification. The difference showed that learners often believed they could check their answers even though their written work did not demonstrate an effective verification method. Their understanding of checking may have been limited to repeating the calculation or reviewing the written numbers for copying errors. Thus, perceived confidence in checking did not necessarily indicate mastery of mathematical verification.

Interpreting information and working without immediate assistance remained within the developing range. Learners appeared willing to begin tasks independently, but this willingness did not always continue once they encountered unfamiliar wording or multistep conditions. Strategy selection obtained a lower measure of -0.24 logits, suggesting uncertainty when several operations or solution paths were possible.

The lowest measures involved continuing after an unsuccessful attempt and explaining a solution. These findings showed that confidence declined when learners had to revise their thinking or make their reasoning visible to others. Some learners appeared comfortable producing a numerical response but uncertain about defending it. Others abandoned the task after their first method failed instead of changing representations or testing another strategy.

Evidence from primary education has shown that mathematics self-efficacy is positively associated with performance, while emotional difficulty can interfere with learners' engagement in mathematical tasks. Živković et al. found that self-efficacy contributed significantly to mathematics performance among fifth-grade learners even when emotional factors were considered. The present pattern supports the view that confidence matters most when it helps learners sustain productive action, rather than when it reflects only a general feeling of being good at mathematics (Živković et al., 2023).

Table 4. *Reasoning-Confidence Calibration Patterns According to Reasoning Level*

Mathematical Reasoning Level	Aligned Confidence, %	Overconfidence, %	Underconfidence, %	Predominant Pattern
Below expected reasoning level	25.40	56.80	17.80	Overconfidence
Approaching expected reasoning level	45.90	34.00	20.10	Aligned confidence
At or above expected reasoning level	55.20	12.90	31.90	Aligned confidence
Overall	42.70	36.10	21.20	Partially aligned

Note. Calibration was based on the standardized confidence measure minus the standardized reasoning measure. Scores from -0.50 to 0.50 were classified as aligned, scores above 0.50 as overconfidence, and scores below -0.50 as underconfidence.

Only 42.70 percent of the learners demonstrated confidence that was reasonably aligned with their measured mathematical reasoning. More than one third were classified as overconfident, while approximately one fifth showed underconfidence. The findings indicated that a considerable portion of the learners did not accurately judge their readiness to solve Grade 6 mathematical problems.

Overconfidence was most pronounced among learners whose mathematical reasoning was below the expected level. In this group, 56.80 percent reported confidence that exceeded their demonstrated reasoning ability. These learners often indicated that they were confident in selecting operations or checking answers, yet their assessment responses contained unsupported choices, incomplete representations, and ineffective verification. Such overconfidence may reduce the likelihood of seeking clarification or carefully reviewing an answer because the learner does not recognize the extent of the reasoning gap.

A different pattern appeared among learners who performed at or above the expected reasoning level. Although most had aligned confidence, 31.90 percent underestimated their own capability. Their written solutions showed appropriate representations, valid strategies, and acceptable explanations, but their confidence responses indicated uncertainty. These learners may possess the required reasoning skills but hesitate to volunteer solutions, explain their work publicly, or attempt unfamiliar tasks without reassurance.

The calibration results showed why confidence and performance should not be interpreted independently. High confidence was not always accompanied by strong reasoning, while low confidence did not always indicate a lack of ability. Liu et al. noted that self-efficacy and achievement may influence one another through prior successes, failures, effort, task difficulty, and learners' interpretation of their own performance. They also emphasized that overestimation can weaken the expected positive relationship between confidence and achievement when confidence judgments are not accurately calibrated (Liu et al., 2024).

Table 5. Robust Regression of Mathematical Reasoning Ability on Problem-Solving Confidence

Predictor	Unstandardized Coefficient, B	Robust SE	Standardized Coefficient, β	z	p	Bootstrap 95% Confidence Interval
Constant	-0.39	0.07		-5.57	< .001	[-0.53, -0.25]
Problem-solving confidence	0.44	0.09	0.41	4.89	< .001	[0.26, 0.61]

Model summary: Robust explained variance = 0.18; residual median absolute deviation = 0.62 logits.
Sensitivity analysis: Spearman's ρ = 0.43, p < .001, bootstrap 95% confidence interval [0.26, 0.58].

The robust regression showed that problem-solving confidence was a significant positive predictor of mathematical reasoning ability. A one-logit increase in confidence corresponded with an estimated increase of 0.44 logits in mathematical reasoning. The standardized coefficient of 0.41 indicated a moderate association. The bootstrap confidence interval did not include zero, confirming that the result remained stable across repeated resampling.

The Spearman rank-order correlation of 0.43 supported the regression result. Learners with stronger confidence generally demonstrated higher reasoning ability, while those with lower confidence tended to obtain lower reasoning measures. Agreement between the robust regression and the nonparametric sensitivity analysis indicated that the relationship was not produced merely by extreme scores or a normality assumption.

Despite its statistical significance, confidence accounted for only 18 percent of the variation in mathematical reasoning. Most of the differences in reasoning remained associated with factors not included in the model, such as conceptual knowledge, reading comprehension, arithmetic fluency, metacognitive monitoring, familiarity with problem structures, and quality of prior learning experiences. Confidence was therefore important but insufficient as a single explanation for reasoning performance.

Öztürk et al. reported that mathematics self-efficacy, reading comprehension, and metacognition were related to mathematical problem-solving performance. Their findings support the present result by showing that confidence operates together with cognitive and self-regulatory processes rather than replacing them (Öztürk et al., 2024).

The cross-sectional design did not establish whether confidence strengthened reasoning, successful reasoning developed confidence, or both processes occurred together. Longitudinal evidence has supported a reciprocal interpretation in which previous achievement can shape later confidence and confidence can affect learners' willingness to persist and engage with subsequent tasks. This interpretation was consistent with the moderate rather than exceptionally strong relationship found in the present analysis (Sakellariou, 2022).

CONCLUSION

The Grade 6 learners demonstrated moderate mathematical reasoning gaps, with the most serious difficulties found in logical justification, answer verification, mathematical representation, and strategy selection. Their problem-solving confidence remained at a developing level and weakened most when they were required to explain their solutions, revise an unsuccessful approach, or work without immediate assistance. Although problem-solving confidence was significantly and positively associated with mathematical reasoning, the calibration analysis showed that many learners either overestimated or underestimated their actual capability. These findings indicate that improving confidence alone is insufficient unless it is accompanied by stronger reasoning, accurate self-evaluation, and opportunities to examine the validity of solutions. It is therefore recommended that mathematics instruction include diagnostic tasks that reveal where reasoning breaks down, structured questioning that requires learners to explain why a strategy is appropriate, multiple representations of mathematical relationships, and routine verification through estimation, inverse operations, substitution, and comparison. Teachers should also provide specific feedback that identifies both successful and incomplete reasoning, allow learners to revise failed solutions, and use confidence-rating activities before and after problem solving to help learners judge their abilities more accurately. Small-group support may be provided to learners with severe reasoning gaps, while learners with adequate performance but low confidence should be given guided opportunities to present and defend solutions. School-based mathematics programs should likewise strengthen teachers' use of reasoning-centered assessment and confidence-calibrating practices, while future studies may examine additional factors such as reading comprehension, arithmetic fluency, metacognitive regulation, and prior achievement through longitudinal or intervention-based research.

References

Di Leo, I., & Muis, K. R. (2020). Confused, now what? A cognitive-emotional strategy training (CEST) intervention for elementary students during mathematics problem solving. *Contemporary Educational Psychology, 62*, Article 101879. doi:10.1016/j.cedpsych.2020.101879

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics, 96*(1), 1–16. doi:10.1007/s10649-017-9761-8

Jonsson, B., Granberg, C., & Lithner, J. (2020). Gaining mathematical understanding: The effects of creative mathematical reasoning and cognitive proficiency. *Frontiers in Psychology, 11*, Article 574366. doi:10.3389/fpsyg.2020.574366

Liu, R., Jong, C., & Fan, M. (2024). Reciprocal relationship between self-efficacy and achievement in mathematics among high school students. *Large-scale Assessments in Education, 12*, Article 14. doi:10.1186/s40536-024-00201-2

Nuutila, K., Tapola, A., Tuominen, H., Molnár, G., & Niemivirta, M. (2021). Mutual relationships between the levels of and changes in interest, self-efficacy, and perceived difficulty during task engagement. *Learning and Individual Differences, 92*, Article 102090. doi:10.1016/j.lindif.2021.102090

Organisation for Economic Co-operation and Development. (2023). *PISA 2022 results, Volume I and II: Country notes: Philippines*. OECD Publishing. doi:10.1787/a0882a2d-en

Öztürk, M., Sarikaya, İ., & Ada Yıldız, K. (2024). Middle school students' problem solving performance: Identifying the factors that influence it. *International Journal of Science and Mathematics Education, 22*(6), 1363–1379. doi:10.1007/s10763-023-10423-5

Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM Mathematics Education, 52*(1), 33–44. doi:10.1007/s11858-019-01118-9

Säfström, A. I., Lithner, J., Palm, T., Palmberg, B., Sidenvall, J., Andersson, C., Boström, E., & Granberg, C. (2024). Developing a diagnostic framework for primary and secondary students' reasoning difficulties during mathematical problem solving. *Educational Studies in Mathematics, 115*(2), 125–149. doi:10.1007/s10649-023-10278-1

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. doi:10.3389/fpsyg.2022.941253

Street, K. E. S., Malmberg, L.-E., & Schukajlow, S. (2024). Students' mathematics self-efficacy: A scoping review. *ZDM Mathematics Education, 56*(2), 265–280. doi:10.1007/s11858-024-01548-0

Street, K. E. S., Malmberg, L.-E., & Stylianides, G. J. (2022). Changes in students' self-efficacy when learning a new topic in mathematics: A micro-longitudinal study. *Educational Studies in Mathematics, 111*(3), 515–541. doi:10.1007/s10649-022-10165-1

United Nations Children's Fund. (2025, December 18). *Latest SEA-PLM 2024 reveals more learners in the Philippines reaching higher proficiency in math and reading, but disparities remain* [Press release]. UNICEF Philippines.

Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Korhonen, J., Aunio, P., & Laine, A. (2024). Task characteristics associated with mathematical word problem-solving performance among elementary school-aged children: A systematic review and meta-analysis. *Educational Psychology Review, 36*(4), Article 117. doi:10.1007/s10648-024-09954-2

Xie, Y., Zeng, F., & Yang, Y. (2024). A meta-analysis of the relationship between metacognition and academic achievement in mathematics: From preschool to university. *Acta Psychologica, 249*, Article 104486. doi:10.1016/j.actpsy.2024.104486

Yang, Y., Maeda, Y., & Gentry, M. (2024). The relationship between mathematics self-efficacy and mathematics achievement: Multilevel analysis with NAEP 2019. *Large-scale Assessments in Education, 12*, Article 16. doi:10.1186/s40536-024-00204-z

Zakariya, Y. F. (2022). Improving students' mathematics self-efficacy: A systematic review of intervention studies. *Frontiers in Psychology, 13*, Article 986622. doi: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*(3), 579–601. doi:10.1007/s11218-023-09760-8