IS CONFIDENCE MORE IMPORTANT THAN KNOWLEDGE?

Alina Kadluba[1], Madeleine Kingan[2], Andreas Obersteiner[1], Jon R. Star[2]

[1]Technical University of Munich, [2]Harvard Graduate School of Education

It is well documented that cognitive and non-cognitive variables are related to students' academic performance. However, their specific predictive power is yet unclear. This study examined how prior performance, confidence, and calibration bias (under- and over-confidence) predict short-term subsequent mathematics performance across Grades 5–8. Students completed tasks of increasing difficulty and indicated confidence for each response. Prior performance consistently predicted later success, while confidence had an increasing effect with schooling. Under-confidence reduced subsequent performance, whereas over-confidence was slightly beneficial in older students. Findings highlight the diagnostic and motivational role of confidence and suggest that fostering self-evaluation and confidence can enhance learning outcomes.

THEORETICAL BACKGROUND

Both cognitive variables (e.g., prior knowledge, general intelligence) and non-cognitive variables (e.g., confidence, motivation, interest) relate to students' academic performance (e.g., Stankov et al., 2012). While cognitive variables as predictors of later student outcomes have been extensively studied (for an overview, see Simonsmeier et al., 2022), non-cognitive variables recently attracted increasing interest in educational research. Confidence—defined as task-specific metacognitive judgments made after solving a task (Efklides, 2011)—emerges as particularly relevant in mathematics achievement (Foster, 2016). Confidence not only contributes to performance but also shapes students' learning behaviors, including persistence, risk-taking, and effort allocation (Stankov et al., 2014).

Confidence is a retrospective, task-specific self-evaluation of performance (Efklides, 2011). It differs from self-efficacy, which is prospective and reflects beliefs about future task success (Bandura & Wessels, 1997), and from academic self-concept, which represents generalized competence perceptions (Stankov et al., 2012). Accurate confidence judgments rely on metacognitive skills, which are essential for effective learning and self-regulation (Stankov et al., 2012).

In mathematics, students often guess when uncertain, and traditional assessment may reinforce this behavior by encouraging students to attempt every task rather than leave responses blank, since even a guessed answer may earn points (Foster, 2016). Confidence judgments provide insight into whether correct answers reflect genuine understanding or guessing, offering diagnostic information for teachers and promoting self-reflection among students (Foster, 2022). Incentivizing accurate judgments, for

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 723-730. PME.

example through reward-penalty systems, can further enhance reflective responding (Gardner-Medwin & Gahan, 2003).

Research demonstrates that confidence is a strong predictor of academic performance, sometimes even exceeding traditional cognitive indicators such as prior achievement (Stankov et al., 2017). While some studies capture confidence and performance on the same tasks (Stankov et al., 2014; Stankov et al., 2017), less is known about how prior confidence predicts subsequent task performance, particularly over short-term sequences.

Confidence judgments are not always accurate. Individuals often display over- or under-confidence, a phenomenon known as *calibration bias* (Koriat, 2012). Lower-performing students tend to be overconfident, whereas higher-performing students may underestimate their competence (Gigerenzer et al., 1991). Accurate self-judgment is important, as misjudgments may lead to inadequate help-seeking or disengagement. Yet, slight over-confidence can enhance motivation, persistence, and resilience (Bouffard & Narciss, 2011).

Confidence is influenced by task- and learner-related factors. Task format affects judgments. Multiple-choice tasks can induce over-confidence due to option familiarity or guessing opportunities (Pallier et al., 2002). Task difficulty also matters. Confidence aligns more closely with performance on easier tasks, whereas difficult tasks increase miscalibration (Lichtenstein et al., 1981). Learner characteristics such as age and prior experience also shape calibration, with developmental studies suggesting that accuracy generally improves with age, though findings are inconsistent for adolescents (e.g., Weil et al., 2013). Previous failures further modulate confidence. Repeated low-confidence experiences can form negative internal signals that influence subsequent task engagement and strategy use (Koriat, 2012).

Research Gaps

Despite growing interest, several questions remain. Most research examines confidence and performance within the same task, leaving the predictive role of prior confidence for subsequent tasks unclear. Similarly, little is known about how calibration bias affects next-task performance or varies with task difficulty and grade level. Understanding these short-term dynamics is particularly relevant for classroom practice, as teachers continuously respond to emerging student difficulties. Insights into these processes can inform interventions targeting not only knowledge acquisition but also the development of metacognitive skills and accurate self-evaluation.

In sum, studying the predictive power of confidence provides both theoretical insight and practical guidance for educational assessment and instruction.

THE PRESENT STUDY

Previous research has highlighted the importance of confidence and calibration bias (i.e., over- and under-confidence) for student learning. However, it remains unclear

how prior confidence, prior performance, and prior calibration bias influence subsequent performance, and how calibration bias varies across task difficulty and grades. To address these gaps, students from grades 5–8 completed tasks at three difficulty levels (easy, medium, hard).

Research Questions and Hypotheses

RQ1a. To what extent do prior performance, prior confidence, and their interaction predict subsequent performance across grade levels?

H1a. Prior confidence is expected to significantly predict subsequent performance, even when controlling for prior performance (Stankov et al., 2014). Further, performance and confidence are hypothesized to interact in predicting subsequent outcomes, with these effects varying across grades due to age-related differences in confidence judgments (Fleming & Lau, 2014; Stankov et al., 2009).

RQ1b. To what extent do under- or over-confidence influence subsequent performance, after controlling for prior performance?

H1b. Under-confidence is expected to negatively impact subsequent performance by reducing engagement, whereas slight over-confidence may foster motivation, persistence, and resilience, thereby positively influencing performance (e.g., Metcalfe, 2009).

RQ2a. Do accurately calibrated responses, under-confidence, and over-confidence differ based on task difficulty?

H2a. Students are expected to show more accurate judgments on easier tasks and greater over-confidence on harder tasks, as individuals who struggle to solve a task are often unaware of their lack of knowledge (Kruger & Dunning, 1999). No specific expectation is formulated for under-confidence.

RQ2b. Do accurately calibrated responses, under-confidence, and over-confidence differ between grade levels?

H2b. Accuracy is expected to improve with grade level. Given inconsistent evidence on developmental changes in metacognitive accuracy (Fleming & Lau, 2014; Foster, 2016, 2022; Stankov et al., 2009; Weil et al., 2013), no clear hypothesis regarding grade-level differences in under- and over-confidence was specified.

This study aims to clarify how cognitive (performance) and non-cognitive (confidence) variables interact to influence achievement and how these relationships vary with task difficulty and schooling.

METHODS

The sample included N = 3,421 students from grades 5–8 across 98 Romanian schools and 93 cities.

Students completed the Romanian Kangaroo mathematics test, which includes 30 multiple-choice tasks (10 per difficulty level: easy, medium, and hard) covering

arithmetic, algebra, geometry, logic, combinatorics, probability, and functions. Tasks were presented in increasing difficulty and the students were asked to complete a minimum of 25 tasks. For each response, students indicated confidence ("sure"), when they were sure about their solution. Students' reported confidence was embedded in a reward–penalty scoring scheme: Correct answers marked as "sure" yielded two additional points, whereas incorrect answers marked as "sure" incurred a point penalty. This design discouraged indiscriminate use of the "sure" option.

Instruments and Measures

Performance: Performance was operationalized as the average proportion of correctly solved items per difficulty level (range: 0–1).

Confidence: Confidence was operationalized as the average proportion of "sure" responses per difficulty level (range: 0–1).

Calibration bias: For each task, responses were scored as −1 (correct but unsure), 0 (correct and sure, or incorrect and unsure), or +1 (incorrect but sure), then converted to proportions for under-confidence, accurate judgment, and over-confidence. These measures are partly related to performance and confidence which should be considered when interpreting the findings. Yet, since the coding depends on item-level performance and confidence, they nonetheless capture additional variance in students' self-judgment and can provide insight into the effects of under-confidence, accurate judgment, and over-confidence on subsequent performance.

Data Analysis

Tasks were presented in order of increasing difficulty: easy, medium, then hard. Linear mixed-effects models reflected this order by predicting medium-task performance from prior easy-task scores, and hard-task performance from prior medium-task scores, with city and school as random effects. Models for calibration bias included prior performance as a covariate. Analyses were conducted separately by grade to examine developmental differences. Differences in calibration across task difficulty were tested with repeated-measures ANOVAs, and between-grade differences with one-way ANOVAs. All analyses were conducted in R.

RESULTS

Descriptive Analysis: Performance and Confidence

Students' performance and confidence varied across task difficulty and grade levels. Overall accuracy was moderate (M = .46), declining from easy (M = .71) to medium (M = .48) and hard tasks (M = .27). Confidence followed a similar pattern, decreasing from easy (M = .57) to medium (M = .37) and hard tasks (M = .25).

Relationship Between Performance and Confidence

At the student level, aggregated performance and confidence were moderately correlated (r = .53), with correlations decreasing for harder tasks and slightly across

grades. At the task level, phi coefficients indicated moderate alignment between correct responses and confidence (overall $\varphi = .35$), declining with task difficulty but increasing across grades ($\varphi = .41$ in Grade 8), suggesting that older students more accurately judged their answers.

Predictors of Subsequent Performance (RQ 1a and 1b)

Across all grades, prior performance was the strongest predictor of subsequent performance. Performance on easy tasks moderately predicted performance on medium tasks ($\beta = .40–.50$), indicating that earlier success reliably transferred to moderately more difficult problems. Likewise, performance on medium tasks predicted performance on hard tasks ($\beta = .22–.35$).

Prior confidence contributed additional explanatory power beyond performance. Confidence on easy tasks positively predicted performance on subsequent medium tasks across all grades ($\beta = .10–.31$), and confidence on medium tasks predicted performance on subsequent hard tasks ($\beta = .07–.14$). These effects, however, were consistently smaller than those of prior performance. Interactions between performance and confidence were observed primarily for more demanding transitions: for the prediction of medium-task performance, interactions were significant only in Grades 5 and 6 ($\beta = .06–.08$), whereas for the prediction of hard-task performance, interactions were significant across all grades and strongest in Grade 8 ($\beta = .20$). This suggests that, especially in higher grades, students benefited most when strong prior knowledge was combined with high confidence.

Calibration bias shaped subsequent performance. Under-confidence consistently predicted lower performance on later tasks across all grades ($\beta = -.09$ to $-.19$), even after controlling for prior performance. Over-confidence showed no systematic effects, except in Grade 8, where it was positively associated with subsequent performance ($\beta = .16$). Overall, prior performance remained the most stable and dominant predictor across all analyses.

Calibration: Accuracy, Under- and Over-Confidence (RQ 2a and 2b)

Task difficulty systematically influenced students' calibration. Accurate judgments were highest on easy tasks (69%), slightly lower on medium and hard tasks (65–66%). Under-confidence decreased slightly with task difficulty (23% to 18%), whereas over-confidence increased markedly (9% to 16%). Developmentally, accurate judgment improved with grade level (Grade 5: 66%, Grade 8: 71%), and under-confidence declined (23% to 19%), while over-confidence remained stable across grades (10–12%). These patterns suggest that students gradually improve in monitoring their performance, particularly in recognizing correct responses, yet tend to increasingly overestimate performance on the hardest tasks.

DISCUSSION

This study examined how cognitive and non-cognitive factors—specifically prior performance, prior confidence, and prior under- or over-confidence—influence subsequent performance in mathematics tasks of varying difficulty across Grades 5–8. Overall, performance and confidence declined with task difficulty. Confidence correlated moderately with performance, with stronger alignment on easier tasks and in higher grades.

Predictive Role of Prior Performance and Confidence

Consistent with prior research (Simonsmeier et al., 2022), prior performance was the strongest predictor of subsequent performance across all grades, suggesting that a solid knowledge foundation supports success in later tasks. In contrast, the predictive effect of prior confidence was smaller in younger grades but increased with age, reaching similar importance to prior performance in Grade 8. This indicates that situational confidence can have an immediate impact on engagement and persistence, even without external feedback. Under-confidence consistently predicted lower subsequent performance, likely reflecting reduced motivation to engage with tasks despite adequate ability (e.g., Metcalfe, 2009), whereas over-confidence had no detrimental effect and even slightly enhanced performance in Grade 8. It should be noted that calibration bias measures depend on prior performance and confidence. For example, a student who answers all items correctly cannot be over-confident, and a student who answers most items incorrectly cannot show high under-confidence. Despite these constraints, the observed associations indicate that the accuracy of students' self-judgments meaningfully influences subsequent performance, as under-confidence, though only possible in students with some correct answers, still predicts lower subsequent performance.

These findings highlight that non-cognitive variables, particularly confidence, influence learning outcomes directly. Students' momentary beliefs about their abilities appear to act as catalysts for task engagement, supporting the idea that fostering positive self-beliefs can enhance performance in real time.

Differences in Calibration Bias Across Tasks and Grades

Calibration accuracy varies with task difficulty and grade level. Accurate judgment and under-confidence decreased with more challenging tasks, while over-confidence increased. Grade 8 students exhibited the highest alignment between confidence and performance, likely due to more advanced metacognitive skills or higher solution rates providing internal cues ("feeling of rightness") for accurate judgments (Thompson et al., 2011). These results suggest that metacognitive abilities and the ability to use confidence strategically develop with age, supporting more effective self-regulated learning in older students.

Task difficulty also influenced confidence calibration. Students were more accurate on easier tasks, consistent with the notion that correct solutions generate clearer internal

cues for confidence judgments. The observed increase in over-confidence on harder tasks aligns with the Dunning-Kruger effect, where limited knowledge impedes accurate self-assessment (Kruger & Dunning, 1999).

Practical Implications

Our findings underscore the importance of incorporating non-cognitive variables into classroom practice. Teachers can ask students to indicate confidence alongside their answers, encouraging reflection on both knowledge and judgment accuracy. Providing feedback on correctness and confidence can strengthen metacognitive monitoring over time. Importantly, this study shows that even without external feedback, students' confidence affected subsequent performance, highlighting the potential benefits of fostering positive self-beliefs. Because confidence becomes particularly influential by Grade 8, interventions targeting situational confidence at this stage may yield the greatest performance benefits. Mild over-confidence in classroom settings may even motivate students to attempt challenging problems.

Limitations

Several limitations should be considered. First, task completion order was not strictly controlled, although most students likely followed the intended sequence. Second, some confidence ratings may not reflect full deliberation, though the reward-penalty system likely improved reliability. Third, multiple-choice formats allowed occasional correct guesses, potentially inflating alignment between performance and confidence. Nevertheless, consistent patterns across tasks and grades support the robustness of our findings.

Conclusion

This study demonstrates that both cognitive and non-cognitive factors shape academic performance. Prior performance reliably predicts subsequent success, while confidence increasingly influences outcomes with age. Under-confidence consistently impairs performance, whereas over-confidence is neutral or beneficial in older students. These findings highlight the importance of supporting not only knowledge acquisition but also confidence development. Integrating confidence ratings and feedback into classroom practice can enhance metacognitive monitoring, motivate engagement, and foster learning outcomes.

References

Bandura, A., & Wessels, S. (1997). *Self-efficacy*. Cambridge University Press Cambridge.

Bouffard, T., & Narciss, S. (2011). Benefits and risks of positive biases in self-evaluation of academic competence: Introduction. *Int. Journal of Educational Research 50*(4), 205-208.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist 46*(1), 6-25.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience 8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Foster, C. (2016). Confidence and competence with mathematical procedures. *Educational Studies in Mathematics*, *91*, 271-288. https://doi.org/10.1007/s10649-015-9660-9

Foster, C. (2022). Implementing confidence assessment in low-stakes, formative mathematics assessments. *Int. Journal of Science and Mathematics Education*, *20*(7), 1411-1429.

Gardner-Medwin, A. R., & Gahan, M. (2003). Formative and summative confidence-based assessment. Proceedings of the 7th CAA Conference, Loughborough.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. *Psychological Review 98*(4), 506.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review 119*(1), 80. https://doi.org/10.1037/a0025648

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology 77*(6), 1121. https://doi.org/10.1037//0022-3514.77.6.1121

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1981). Calibration of probabilities: The state of the art to 1980.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science 18*(3), 159-163. https://doi.org/10.1111/j.1467-8721.2009.01628.x

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *The Journal of General Psychology 129*(3), 257-299.

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Ed. Psychologist 57*(1), 31-54.

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences 22*(6), 747-758. https://doi.org/10.1016/j.lindif.2012.05.013

Stankov, L., Lee, J., Morony, S., Lee, Y. P., Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., Abduljabbar, A. S., & Seaton, M. (2014). Noncognitive psychological processes and academic achievement *Educational Psychology*, *34*(1).

Stankov, L., Lee, J., & Paek, I. (2009). Realism of confidence judgments. *European Journal of Psychological Assessment*, *25*(2), 123-130. https://doi.org/10.1027/1015-5759.25.2.123

Stankov, L., Morony, S., & Lee, Y. P. (2017). Confidence: the best non-cognitive predictor of academic achievement? In *Noncognitive Psychological Processes and Academic Achievement* (pp. 19-38). Routledge. https://doi.org/10.1080/01443410.2013.814194

Thompson, V. A., Turner, J. A. P., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology 63*(3), 107-140.

Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., Dolan, R. J., & Blakemore, S.-J. (2013). The development of metacognitive ability in adolescence. *Consciousness and Cognition 22*(1), 264-271.