EXPLORING FACTORS INFLUENCING SUCCESS IN MATHEMATICAL PROBLEM SOLVING

EXPLORANDO LOS FACTORES QUE INFLUYEN EN EL ÉXITO EN LA RESOLUCIÓN DE PROBLEMAS MATEMÁTICOS

Sam Rhodes
Georgia Southern University
srhodes@georgiasouthern.edu

Rick Bryck
Landmark College
rickbryck@landmark.edu

Antonio Gutierrez de Blume
Georgia Southern University
agutierrez@georgiasouthern.edu

Supporting students in becoming effective problem solvers is a critical component of K-12 mathematics instruction. Unfortunately, little is known about the factors that are related to problem solving proficiency in middle school students. We report the results of a study that employed a hierarchical linear regression analysis to examine the factors that influenced the problem-solving scores in a sample of 213 sixth and seventh grade students. Results support theoretical predictions that metacognition, executive function, student beliefs, and content knowledge all relate to problem solving proficiency.

Keywords: Problem solving; Metacognition; Affect, Emotion, Beliefs, and Attitudes.

Research has suggested that teaching with and through mathematical problem solving (PS) has vast potential for engaging all students in mathematical practices that support increased motivation, achievement, positive beliefs about mathematics, and conceptual understanding of mathematical concepts (Boaler, 2002; Boaler & Staples, 2008; Cai, 2003; Lester & Cai, 2013; Stein et al., 2003). Consequently, PS is often considered to be a focal point of effective mathematics instruction (NCTM, 2014; 2018). However, relatively little is known about how to effectively teach mathematical PS (Lester & Cai, 2016).

Proficiency in PS is thought to be related to interactions between multiple factors such as metacognition (MC), strategic thinking, content knowledge, and the beliefs and dispositions that students have about themselves and mathematics (Chapman, 2015; Schoenfeld, 2013). Previous studies have sought to investigate the relationship between individual factors and PS, with factors such as MC (e.g., Özsoy & Ataman, 2009; Tan & Limjap, 2018), executive function (EF; Viterbori et al., 2017), and student anxieties and beliefs (Kramarski et al., 2010; Mayer & Wittrock, 2006) being investigated. However, few studies to date have empirically explored the relationship between PS and a combination of cognitive, metacognitive, and affective factors. The current study–which was part of a larger study on EF and mathematical PS, funded by the Advanced Education Research and Development Fund (AERDF)–responds to this need by using a hierarchical linear regression to determine which factors explained unique variance in the PS scores of middle school students. Better understanding these relationships may be vital in developing future PS interventions and instruction to support students.

Review of Related Literature

Mathematical Problem Solving and Executive Function

Executive function refers to goal-directed, attention regulation skills that predict a wide range of important academic outcomes, including mathematics performance (Zelazo et al., 2016). These include the ability to keep information in mind (working memory; WM), the ability to stop oneself from responding immediately (inhibition), and the ability to think in multiple ways or perspectives (cognitive flexibility). Theoretical and empirical evidence show a relationship between these three core components of EF and mathematics: WM (Raghubar et al., 2010), inhibition (Lemaire & Lecacheur, 2011), and cognitive flexibility (Yeniad et al., 2013).

Moreover, EF skills appear to be particularly needed for multi-step PS skills, and evidence suggests that developing EF skills support the development of math PS, and vice versa (Clements et al., 2016). WM, in particular, has been shown to be an important predictor of math PS accuracy (Passolunghi & Siegel, 2001; Swanson et al., 2008; Steinberg & Roditi, 2018). Inhibition, in turn, is thought to be critical in suppressing inefficient strategies and ignoring irrelevant information during PS (Bull & Lee, 2014), and cognitive flexibility allows students to fluidly shift between mathematical representations. These three core EF components seem to fill unique needs at different stages of the PS process (Viterbori et al., 2017).

Mathematical Problem Solving and Metacognition

Research has suggested that metacognitive knowledge—one's awareness of their own cognition—and regulation—the ability to control one' cognition—are critical components of PS (Favell, 1976, 1979; Lester, 2013; Schoenfeld, 1992, 2013; Stillman & Galbraith, 1998). Specifically, MC has been found to be critical throughout the PS process, impacting everything from how one engages in sense-making, to selecting and utilizing strategies, to implementing plans and reflecting on solutions (Tan and Limjap, 2018). This link has been empirically tested in numerous studies, with findings often suggesting that metacognitive training and supports result in significant improvement in students' PS (e.g., Hensberry & Jacobbe, 2012, Montague et al., 2011, Özsoy & Ataman, 2009).

Moreover, while the use of heuristics or strategic thinking has long been considered a key component of successful PS, as noted above, research has suggested that heuristics are insufficient unless they are coupled with effective metacognitive awareness and regulation. These are necessary to support problem solvers in adapting heuristics to new problems and monitoring what they are doing and evaluating their plan throughout the PS process (Lesh & Zawojewski, 2007; Jitendra et al., 2015). Indeed, students' metacognitive verbalizations have been found to increase along with problem difficulty (Rosenzweig et al., 2011), and students' metacognitive awareness has been found to be especially critical when students utilize new strategies or engage with non-routine problems (Carr & Jessup, 1995).

Mathematical Problem Solving and Affect

Numerous affective factors are thought to impact PS proficiency. These include self-efficacy, beliefs, anxiety, and math identity (Chapman, 2015; Irhamna et al., 2020; Tzohar-Tozen & Kramarski, 2014). Exemplifying these links, WM is used to attend to various functions, meaning that increased anxiety may decrease cognitive performance on other tasks (Moran, 2016). Moreover, student beliefs can support or inhibit students' ability to regulate their cognition, as well as their motivation and persistence as they encounter challenges inherent in PS (Goldin et al., 2016; Tzohar-Tozen & Kramarski, 2014).

Theoretical Framework

Research has suggested that self-regulated learning (SRL) is a key aspect of mathematical PS (Tzohar-Rozen & Kramarski, 2014). SRL theory refers to the ability to control one's learning environment and is posited to encompass cognition, MC, and motivation processes. Several theoretical accounts of SRL have been proposed in the literature (see Panadero, 2017, for a review). For instance, Zimmerman's Cyclical Phases Model (CPM; Zimmerman & Moylan, 2009) describes SRL as a cyclical process involving three parts: (1) forethought (e.g., goal setting, strategic planning, self-efficacy beliefs, and intrinsic motivation); (2) performance and volitional control (e.g., attention focusing, self-instruction, and self-monitoring); and (3) self-reflection (e.g., self-evaluation, attributions, and self-reactions). Similarly, Winne and Hadwin (1998) developed a Metacognitive Perspective Model (MPM) of SRL in which metacognitive processes play a central role. According to the tenets of this model, learners are perceived as being active and involved self-regulated individuals who control their own learning through the implementation of metacognitive monitoring and strategy use, which are central to the goals of the present study. The model was subsequently expanded to include self-regulatory actions (Winne & Hadwin, 2008). Along a similar vein, Efklides (2011) devised the Metacognitive and Affective Model of Self-Regulated Learning (MASRL) in which metacognitive and motivational processes are also key. Even though all these models vary regarding labels and what aspects to include, they all agree that learning is regulated by a variety of dynamic interacting and cyclical cognitive, metacognitive, and motivational factors (Butler & Winne, 1995; Panadero, 2017). The present investigation was centered on the MASRL (Efklides, 2011) and MPM (Winne & Hadwin, 2008), as those are the models in which metacognitive skills play a central role. Specifically, SRL informed both the factors under consideration within the present study and the method of model-building used for the hierarchical linear regression.

Methods

Participants

Participants were drawn from three middle schools from a large, suburban school district located on the West Coast of the United States. All grade 6-8 mathematics teachers at these schools, and their respective students, were given the opportunity to participate in the study. Within this, 115 sixth grade students and 98 seventh grade students completed all measures, yielding a total sample of 213 for the analyses reported herein. The participants identified their gender as male (40.8%), female (53.1%), non-binary (1.9%), or other (2.8%), with 1.4% of participants electing not to specify a gender. The participants also identified as Hispanic/Latinx (29.6%), Middle Eastern (28.6%), 2 or more races (16.4%), Asian (6.6%), Black/African American (4.2%), or White (2.8%), with 11.7% of participants preferring not to specify.

Research Instruments and Scoring

Executive Function. Students' EF was measured using the Adaptive Cognitive Evaluation (ACE; Younger et al., 2022)—gamified versions of well-known computerized cognitive tasks of three core EFs: inhibition (Flanker), working memory (change detection), and cognitive flexibility (task switching; Miyake et al., 2000). Mean and standard deviation scores were calculated for both reaction time and accuracy measures. Difference scores were the primary variables of interest for each measure, including "Flanker effect" (incongruent minus congruent trials), "K" (estimate of WM capacity, hits minus false alarms), "WM filtering" (set size 2 with 2 distractors minus set size 2), and "switch cost" (task switching switch minus stay trials).

Content Knowledge. Given that the partnering district utilized the i-Ready diagnostic assessments (Curriculum Associates, 2022) to gain regular data on students' content knowledge in mathematics, these scores were utilized as a proxy for mathematics content knowledge.

Metacognition. Subjective metacognitive awareness was measured using a shortened version of the Junior Metacognitive Awareness Inventory (MAI, Jr.; Gutierrez de Blume et al., under review; Sperling et al., 2002). The survey was administered using a sliding scale of 1 to 100 to allow for more continuous data collection. In addition, objective MC was measured through confidence judgements in which students predicted how well they would do on the PS measure (see below) prior to taking the measure, and then postdicted their performance immediately after completing the measure. Scores for objective MC were calculated by computing the absolute value of the difference scores between an individual's confidence judgements and their actual performance.

Affective Instruments. Several surveys were used to measure various affective factors. These included the modified Abbreviated Math Anxiety Scale (mAMAS; Carey et al., 2017), and belief scales 1, 5, and 6 of the Indiana Mathematics Beliefs Scales (IMBS; Kloosterman & Stage, 1992), which measured students' beliefs about whether they can solve time-consuming problems, about whether effort increases ability, and about the usefulness of mathematics in their lives, respectively. In addition, students were given five questions about their feelings about mathematics within their classrooms (e.g., "I feel encouraged to solve challenging problems in math") and a question about how close they felt to the subject of mathematics.

All surveys were administered as originally developed, with the following exceptions: 1) all surveys were administered using a sliding scale of 1 to 100; 2) the mAMAS was slightly adapted for American English (e.g., "maths" to "math"); and 3) an abbreviated version of the IMBS (Rhodes et al., forthcoming), was employed to reduce testing fatigue. For each survey, all questions within each scale were averaged to obtain a single score for each distinct construct.

Problem Solving Measure. The researchers developed a distinct 3-item PS measure for each grade, with one item aligning to content taught in each trimester per the pacing guide that was provided by the partnering school district. All measure items were drawn from problems developed by Illustrative Mathematics (IM) and were chosen for inclusion based on A) the degree to which they were cognitively demanding and rigorous; B) the degree to which they aligned to standards that the district identified as priorities within their respective grade levels; and C) the degree to which they required students to show or explain their thinking.

If questions met criteria A and B but not C, problems were slightly modified (e.g., adding directions requiring students to show or explain their thinking). Each problem was scored as being correct or incorrect (referred to herein as accuracy) using answer keys developed by IM. In addition, problems were scored using rubrics developed by an external researcher that assessed the degree to which a student's work demonstrated relevant mathematical understandings, regardless of the correctness of the final answer (referred to herein as understanding). Given the links between the use of heuristics and MC noted above, we intentionally disregarded scoring students' strategies, focusing instead only on the mathematically understandings they demonstrated. This choice was made to avoid scoring in ways that specifically considered MC prior to then correlating the scores with measures of MC. Interrater agreement on scoring for accuracy and understanding was measured by Fleiss' kappa. The Fleiss' kappas were .961 and .880 for accuracy and .703 and .842 for understanding for 6th and 7th grade, respectively.

Analyses

Prior to data analyses, data were screened for requisite statistical assumptions including normality, linearity, and multicollinearity. Data met all statistical assumptions. However, two outliers were detected in the data that were determined to be two standard deviations from the line of best fit. These outliers were removed and omitted from subsequent analyses. To address the research objective, two hierarchical linear regressions were performed: one with PS accuracy as criterion, and a second with PS understanding as criterion. Guided by SRL theory, predictors for each criterion variable were entered in the following order. Cognitive measures (i.e., EF measures, iReady scores) were entered in the first block. Affective factors (e.g., the IMBS) were entered in the second block, and MC factors (e.g., objective metacognitive monitoring accuracy, MAI, Jr. scores) were entered into the third block. Given that all assumptions were met, the results were analyzed, and findings are reported below.

Results

Accuracy

The omnibus model for accuracy as the criterion was statistically significant, $F(4,206) = 47.44$, $p < .001$, $R^2 = .480$. EF were entered as the first block of predictors, and they contributed significant incremental variance to the prediction of accuracy, $\Delta F(1,209) = 8.42$, $\Delta p = .004$, $\Delta R^2 = .039$. Task switching was the only significant predictor of accuracy. Mathematics beliefs, as measured by the IMBS, was entered as the second block of predictors, and it also contributed significant incremental variance to the prediction of accuracy, $\Delta F(1,208) = 13.46$, $\Delta p < .001$, $\Delta R^2 = .058$. Scale 1, measuring the extent to which a student can solve time consuming math problems, was the only significant predictor. Metacognitive monitoring, as assessed by students prediction feeling of knowing judgments, was entered as the third block predictor, and it contributed significant incremental variance to the prediction of accuracy, $\Delta F(1,207) = 141.17$, $\Delta p < .001$, $\Delta R^2 = .366$. Finally, the degree to which students felt encouraged to solve challenging problem in math class, was entered as the fourth block predictor, and it also contributed significant incremental variance to the prediction of accuracy, $\Delta F(1,206) = 6.45$, $\Delta p < .012$, $\Delta R^2 = .016$. Table 1 includes the model coefficients for all four blocks.

Table 1: Regression Coefficients Results for Math Problem Solving Accuracy

Predictor	*b*	*β*	*t*	*p*	$CI_{95\%}$	
					LB	UB
Block 1: Executive Function						
Task Switching	.001	.197	2.90	.004	.000	.002
Block 2: IMBS						
Solving Time Consuming Problems	.007	.196	2.97	.003	.003	.011
Block 3: Objective Metacognitive Monitoring Accuracy						
Prediction Judgments	-.017	-.614	-11.88	< .001	-.020	-.014
Block 4: Student Feelings About Their Mathematics Classroom						
Solving Challenging Math Problems	.003	.142	2.54	.012	.001	.006

Key: *b* = unstandardized regression coefficient; *β* = standardized regression coefficient; 95% confidence interval for the unstandardized regression coefficients; LB = Lower bound value; UB = Upper bound value.
N = 213

The omnibus model for understanding as the criterion was statistically significant, $F(6,204) = 11.46$, $p < .001$, $R^2 = .251$. iReady math achievement scores for Fall 2021 was entered as a block 1 predictor, and it contributed significant incremental variance to the prediction of understanding, $\Delta F(1,209) = 16.46$, $\Delta p < .001$, $\Delta R^2 = .073$. EF were entered as the second block of predictors, and they contributed significant incremental variance to the prediction of accuracy, $\Delta F(1,208) = 4.25$, $\Delta p = .041$, $\Delta R^2 = .019$. Task switching, as a measure of cognitive flexibility, was the only significant predictor of understanding. Mathematics beliefs, as measured by the IMBS, was entered as the third block of predictors, and it also contributed significant incremental variance to the prediction of understanding, $\Delta F(1,207) = 7.85$, $\Delta p < .006$, $\Delta R^2 = .033$. Scale 1, measuring the extent to which a student can solve time consuming math problems, was the only significant predictor. Metacognitive monitoring, as assessed by students prediction feeling of knowing judgments, was entered as the fourth block predictor, and it contributed significant incremental variance to the prediction of understanding, **Δ**$F(1,206) = 19.09$, **Δ**$p < .001$, **Δ**$R^2 = .074$. Postdiction, as another metacognitive monitoring measure (after students have seen the task), was entered as the fifth block predictor, and it contributed significant incremental variance to the prediction of understanding, **Δ**$F(1,205) = 10.26$, **Δ**$p = .002$, **Δ**$R^2 = .038$. Finally, the degree to which students felt encouraged to solve challenging problem in math class, was entered as the sixth block predictor, and it also contributed significant incremental variance to the prediction of understanding, **Δ**$F(1,204) = 4.05$, **Δ**$p = .045$, **Δ**$R^2 = .015$. Table 2 includes the model coefficients for all six blocks.

Table 2: Regression Coefficients Results for Math Problem Solving Understanding

Predictor	*b*	*β*	*t*	*p*	CI95%	
	LB	UB				
	Block 1: Math Achievement					
iReady Math	.004	.270	4.06	< .001	.002	.005
	Block 2: Executive Function					
Task Switching	.000	.136	2.06	.041	.000	.001
	Block 3: IMBS					
Solving Time Consuming Problems	.003	.186	2.80	.006	.001	.005
	Block 4: Objective Metacognitive Monitoring Accuracy					
Prediction Judgments	-.005	-.281	-4.37	< .001	-.007	-.003
	Block 5: Objective Metacognitive Monitoring Accuracy					
Postdiction Judgements	.004	.226	3.20	.002	.001	.006
	Block 6: Student Feelings About Their Mathematics Classroom					
Solving Challenging Math Problems	.002	.136	2.01	.045	.000	.003

Key: *b* = unstandardized regression coefficient; *β* = standardized regression coefficient; 95% confidence interval for the unstandardized regression coefficients; LB = Lower bound value; UB = Upper bound value.
N = 213

Discussion

The purpose of the present investigation was to determine the predictive effect of EF, students' beliefs about their math skills, students' self-report math anxiety, math achievement, and objective metacognitive monitoring on students' PS accuracy and PS understanding. The findings of the study substantiate prior research that demonstrated factors such as anxiety and self-confidence (Irhamna et al., 2020), MC (Lesh & Zawojewski, 2007; Jitendra et al., 2015), content knowledge (Chapman, 2015), and EF (Clements et al., 2016; Viterbori et al., 2017), relate to PS proficiency. This study also extends this research by confirming that cognitive, metacognitive, and affective factors each explain unique variance within the PS scores of middle school students.

Regarding PS accuracy, results revealed that task switching, students' beliefs that they can solve time-consuming math problems, prediction judgments, and students' feelings regarding teachers' encouragement to solve challenging math problems, significantly predicted PS accuracy. Regarding PS understanding, math achievement, task switching, students' beliefs that they can solve time-consuming math problems, prediction and postdiction feeling of knowing metacognitive judgments, and students' feelings regarding teachers' encouragement to solve challenging math problems, were significant.

Cognitively, content knowledge was only predictive within the model of understanding, but not accuracy. Given that the individual correlation between content knowledge and PS understanding, $r = .058$, $p > .05$, was also non-significant, this potentially suggests that correct answers alone are insufficient to consider when examining students' mathematical understandings within PS solutions. EF, however, explained unique variance within both accuracy and understanding. Although task switching, generally considered a proxy of cognitive flexibility, predicted PS accuracy and understanding, the direction was opposite of our theoretical prediction. Our measure of interest, reaction time (RT) "switch cost," is the difference in RT between trials in which the task switches compared to repeat trials. In short, a robust finding is that "switch" trials have longer RTs on average than "repeat" trials, with the longer time (or difference between switch and stay) indicative of the additional EF processes that must occur. Thus, lower switch costs are thought to represent more efficient EF processing; as such, we predicted lower switch costs to be correlated with higher PS performance, but we observed the opposite. It is possible that a complex speed-accuracy trade-off is occurring, such that students that take more time during PS to engage in reflective processing perform similarly during task-switching, that is they pause to reflect after a task switch, which would in turn inflate their observed switch costs.

Interestingly, MC has a rather nuanced relation with students' PS accuracy and understanding. Whereas prediction judgments significantly predicted PS accuracy, both prediction and postdiction judgements predicted PS understanding. This is interesting because the literature on MC has consistently demonstrated that students tend to be more accurate in their postdictions (after the task) than their predictions (e.g., Gutierrez & Schraw, 2015; Hacker et al., 2008), especially when the postdictions are delayed (Thiede et al., 2012), presumably because students' have had additional opportunities to reflect on their performance and adjust their confidence in their performance to more closely align to their actual performance (Gutierrez de Blume et al., 2021).

Further, predictions, which are made before the task, were the best predictors of PS accuracy, uniquely explaining the majority of variability. Conversely, the combined predictive effect of predictions and postdictions on PS understanding was significantly attenuated. It is plausible that PS accuracy, which utilizes similar cognitive processes to metacomprehension judgments, is more closely associated with metacognitive skills than PS understanding. Indeed, this relation between metacomprehension and PS accuracy has been found in domains other than math (e.g., driving ability, Ackerman et al., 2010; physics, Sharma & Bewes, 2011).

In the affective domain, students' self-efficacy regarding their ability to solve time consuming mathematics problems was predictive in both models, along with the question measuring the degree to which students felt encouraged to solve challenging problems in their mathematics classes. Interestingly, mathematics anxiety, student beliefs about whether effort increases ability, and students' perceptions of how close they felt to mathematics were not unique predictors within either model, even though their individual correlations were significant with both accuracy and understanding. This potentially suggests that the variance was subsumed by other factors such as MC, EF, and self-efficacy.

Limitations and Avenues for Future Research

There are number of limitations to consider from the present study. Chiefly, all factors were measure at single time points. However, many of these factors are dynamic and can change based on unique situations, contexts, and environment. Given that correlational analyses are limited in their ability to attend to factors that are context dependent (Schoenfeld, 2002), future research should consider utilizing state-based measures along with qualitative analyses to better understand how these factors interact in-the-moment, and how best to support students in becoming proficient problem solvers.

Secondly, although the ACE offers the advantage of being short and gamified (improving engagement), these shortened versions of classic cognitive batteries, typically involving hundreds of trials per condition, likely produce more variability and noise. Moreover, these computerized EF measures assess "core" EF, but may not fully predict "real-world" EF, such as those required in a math classroom. Finally, the present study was limited in the affective factors considered and thus future research should explore additional factors (e.g., motivation) and more robust measures of students' math identity.

Implications and Recommendations for Research, Theory, and Practice and Conclusion

As utilizing rigorous PS within instruction continues to be a focal point of K-12 mathematics courses, the results of the present study offer critical insights into factors related to PS proficiency that are essential to consider in supporting student learning. As previously suggested in research (e.g., Chapman, 2015), PS proficiency is complex, comprised of myriad cognitive, metacognitive, and affective factors. These results suggest that future research and instruction on PS should simultaneously attend to cognitive, metacognitive, and affective factors within PS, rather than parsing them out for individual attention. Thus, strengthening student PS may best be approached via a holistic, multi-pronged instructional approach that incorporates each of these factors, rather than attempting to teach these skills independently.

Moreover, the present investigation demonstrated that an objective measure of MC (monitoring accuracy before and after a task) was the best predictor of PS accuracy, and a significant predictor of PS understanding, albeit to a lesser extent. This has relevant implications for further research targeting these factors more deeply, and for the theories of SRL, MC and EF. For educators, this supports previous findings on the importance of attending to MC within PS instruction (e.g., Montague et al., 2011), and highlights the need for continued attention to developing students' metacognition within the context of instruction on mathematical PS.

Acknowledgements

The research reported here was supported by the EF+Math Program of the Advanced Education Research and Development Fund (AERDF) through funds provided to CueThinkEF+. The perspectives expressed are those of the authors and do not represent views of the EF+Math Program or AERDF.

References

Ackerman, M., Crowe, M., Vance, D., Wadley, V., Owsley, C., & Ball, K. (2010). The impact of feedback on self-rated driving ability and driving self-regulation among older adults. *The Gerontologist, 51*(3), 367-378. https://doi.org/10.1093/geront/gnq082

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning.* Routledge.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record, 110*(3), 608–645. https://doi.org/10.1177/016146810811000

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8*(1), 36–41. https://doi.org/10.1111/cdep.12059
Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281. https://doi.org/10.3102/00346543065003245
Cai, J. (2003). What research tells us about teaching mathematics through problem solving. In F. K. Lester & R. I. Charles (Eds.), *Research and issues in teaching mathematics through problem solving* (pp. 241–254). National Council of Teachers of Mathematics.
Carey, E. (2017). The modified abbreviated math anxiety scale: A valid and reliable instrument for use with children. *Frontiers in Psychology, 8*, 13. https://doi.org/10.3389/fpsyg.2017.00011
Carr, M., & Jessup, D. L. (1995). Cognitive and metacognitive predictors of mathematics strategy use. *Learning and Individual Differences, 7(3),* 235–247.
Chapman, O. (2015). Mathematics teachers' knowledge for teaching problem solving. *LUMAT: International Journal on Math, Science and Technology Education*, *3*(1), 19–36. https://doi.org/10.31129/lumat.v3i1.1049
Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education, 38,* 136–163. DOI: 10.2307/30034954
Clements, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly*, 36, 79–90. https://doi.org/10.1016/j.ecresq.2015.12.009
Curriculum Associates (2022). Research base for i-*Ready Personalized Instruction for Mathematics:* Executive Summary. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-research-foundations-mathematics-executive_summary-2020.pdf
Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*, 6–25. https://doi.org/10.1080/00461520.2011.538645
Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Erlbaum.
Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.
Goldin, G. A., Hannula, M. S., Heyd-Metzuyanim, E., Jansen, A., Kaasila, R., Lutovac, S., ... Zhang, Q. (2016). Attitudes, beliefs, motivation and identity in mathematics education: An overview of the field and future directions. In *ICME-13 Topical Surveys* (pp. 1–35). https://doi.org/https://doi.org/10.1007/978-3-319-32811-9_1
Gutierrez, A. P., & Schraw, G. (2015). Effects of strategy training and incentives on students' performance, confidence, and calibration. *The Journal of Experimental Education, 83*, 386-404. https://doi.org/10.1080/00220973.2014.907230
Gutierrez de Blume, A. P., Schraw, G., Kuch, F., & Richmond, A. S. (2021). General accuracy and general error factors in metacognitive monitoring and the role of time in predicting metacognitive judgments. CES Psicología, 14(2), 179-208. https://doi.org/10.21615/cesp.5494
Gutierrez de Blume, A.P., Rhodes, S., & Bryck, R. (under review). *Metacognitive awareness amount middle school adolescents: Development and validation of a shortened version of the MAI, Jr.*
Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition Learning, 3*, 101-121. https://doi.org/10.1007/s11409-008-9021-5
Hensberry, K. K. R., & Jacobbe, T. (2012). The effects of Polya's heuristic and diary writing on children's problem solving. Mathematics Education Research Journal, 24(1), 59–85. DOI: 10.1007/s13394-012-0034-7
Irhamna, Amry, Z., Syahputra, H. (2020). Contribution of mathematical anxiety, learning motivation and self-confidence to student's mathematical problem solving. *Budapest International Research and Critics in Linguistics and Education, 3*, 1759-1772. https://doi.org/10.33258/birle.v3i4.1343.g1806
Jitendra, A. K., Petersen-Brown, S., Lein, A. E., Zaslofsky, A. F., Kunkel, A. K., Jung, P. G., & Egan, A. M. (2015). Teaching mathematical word problem solving: The quality of evidence for strategy instruction priming the problem structure. *Journal of Learning Disabilities, 48*(1), 51–72. https://doi.org/10.1177/0022219413487408
Kloosterman, P., & Stage, F. K. (1992). Measuring beliefs about mathematical problem solving. *School Science and Mathematics, 92(3)*, 109–115. https://doi.org/10.1111/j.1949-8594.1992.tb12154.x
Kramarski, B., Weisse, I., & Kololshi-Minsker, I. (2010). How can self-regulated learning support the problem solving of third-grade students with mathematics anxiety? *ZDM - International Journal on Mathematics Education, 42(2)*, 179–193. https://doi.org/10.1007/s11858-009-0202-8

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development, 26*, 282–294. https://doi.org/10.1016/j.cogdev.2011.01.002
Lesh, R., & Zawojewski, J. S. (2007). Problem solving and modeling. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 763–804). Information Age.
Lester, F. K. (2013). Thoughts About Research On Mathematical Problem- Solving Instruction. *The Mathematics Enthusiast, 10(1),*245-278. https://doi.org/10.54870/1551-3440.1267
Lester, F. K., & Cai, J. (2016). Can mathematical problem solving be taught? Preliminary answers from 30 years of research. In P. Felmer, E Pehkonen, & J. Kilpatrick (Eds.), *Posing and solving mathematical problems* (pp. 117–135). https://doi.org/10.1007/978-3-319-28023-3
Mayer, R. E., & Wittrock, R. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed.), (pp. 287-304). Mahwah, NJ: Erlbaum.
Montague, M., Enders, C., & Dietz, S. (2011). Effects of cognitive strategy instruction on math problem solving of middle school students with learning disabilities. *Learning Disability Quarterly, 34*(4), 262–272. https://doi.org/10.1177/0731948711421762
Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin, 142*(8), 831–864. https://doi.org/10.1037/bul0000051
National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. National Council of Teachers of Mathematics
National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations*. Author.
Özsoy, G., & Ataman, A. (2009). The effect of metacognitive strategy training on mathematical problem solving achievement. *International Electronic Journal of Elementary Education, 1(2)*, 68–82.
Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8,* 1-28. https://doi.org/10.3389/fpsyg.2017.00422
Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80(1)*, 44–57. https://doi.org/10.1006/jecp.2000.2626
Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20(2)*, 110–122. https://doi.org/10/fs4ghp
Rhodes, S., Gutierrez de Blume, A., & Bryck, R. (in preparation). *Development and validation of a shortened version of the Indiana Mathematics Beliefs Scales.*
Rosenzweig, C., Krawec, J., & Montague, M. (2011). Metacognitive strategy use of eighth-grade students with and without learning disabilities during mathematical problem solving: A think-aloud analysis. *Journal of Learning Disabilities, 44*, 508– 520. https://doi.org/10.1177/0022219410378445
Schoenfeld, A. H. (1992). Learning to think mathematically: Sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334–370). MacMillan.
Schoenfeld, A.H. (2002). How can we examine the connections between teachers' world views and their educational practices? *Issues in Education, 8*(2), 217-227.
Schoenfeld, A. H. (2013). Reflections on problem solving theory and practice. *The Mathematics Enthusiast, 10(2)*, 9-34. https://doi.org/10.54870/1551-3440.1258
Sharma, M. D., & Bewes, J. (2011). Self-monitoring: Confidence, academic achievement and gender differences in physics. *Journal of Learning Design, 4*(3), 1-13. doi: http://dx.doi.org/10.5204/jld.v4i3.76
Sperling, R. A., Howard, B. C., & Miller, L. A. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51-79. https://doi.org/10.1006/ceps.2001.1091
Stein, M., Boaler, J., & Silver, E. A. (2003). Teaching mathematics through problem solving: Research perspectives. *In Teaching mathematics through problem solving: Grades 6-12*. National Council of Teachers of Mathematics.
Steinberg, J., & Roditi, B. N. (2018). The strategic math classroom: How executive function impacts math learning. *In Executive Function in Education: From Theory to Practice*, 2nd ed (pp. 300–325). Guilford Press.
Stillman, G. A., & Galbraith, P. L. (1998). Applying mathematics with real world connections: Metacognitive characteristics of secondary students. *Educational Studies in Mathematics, 36*, 157–189.
Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*, 343–379. https://doi.org/10.1037/0022-0663.100.2.343

Tan, D. A., & Limjap, A. A. (2018). Filipino students use of metacognitive skills in mathematical problem solving: An emergent model. *International Journal of Development Research, 8*, 20430–20439.
Thiede, K. W., Redford, J. S., Wiley, J., & Griffin, T. D. (2012). Elementary school experience with comprehension testing may influence metacomprehension accuracy among seventh and eighth graders. *Journal of Educational Psychology, 104*(3), 554–564. https://doi.org/10.1037/a0028660
Tzohar-Rozen, M., & Kramarski, B. (2014). Metacognition, motivation, and emotions: Contributions on self-regulated learning to solve mathematical problems. *Global Education Review, 1*(4), 76–94.
Viterbori, P., Traverso, L., & Usai, M. C. (2017). The role of executive function in arithmetic problem-solving processes: A study of third graders. *Journal of Cognition and Development, 18(5)*, 595–616. https://doi.org/10.1080/15248372.2017.1392307
Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.
Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Lawrence Erlbaum Associates.
Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004
Younger, J., O'Laughlin, K., Anguera, J. A., Bunge, S., Ferrer, E. E., Hoeft, F., McCandliss, B., Mishra, J., Rosenberg-Lee, M., Gazzaley, A., & Uncapher, M. (2022). *Better together: Novel methods for measuring and modeling development of executive function diversity while accounting for unity.* PsyArXiv. https://doi.org/10.31234/osf.io/ctzwx
Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2017). *Executive function: Implications for education (NCER 2017-2000)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the Institute website at http://ies.ed.gov/ncer/pubs/20172000/
Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299-315). Routledge. https://doi.org/10.4324/9780203876428