

# Inducing mental set constrains procedural flexibility and conceptual understanding in mathematics

Marci S. DeCaro<sup>1,2</sup>

© Psychonomic Society, Inc. 2016

**Abstract** An important goal in mathematics is to flexibly use and apply multiple, efficient procedures to solve problems and to understand why these procedures work. One factor that may limit individuals' ability to notice and flexibly apply strategies is the mental set induced by the problem context. Undergraduate ( $N = 41$ , Experiment 1) and fifth- and sixth-grade students ( $N = 87$ , Experiment 2) solved mathematical equivalence problems in one of two set-inducing conditions. Participants in the complex-first condition solved problems without a repeated addend on both sides of the equal sign (e.g.,  $7 + 5 + 9 = 3 + \_$ ), which required multistep strategies. Then these students solved problems with a repeated addend (e.g.,  $7 + 5 + 9 = 7 + \_$ ), for which a shortcut strategy could be readily used (i.e., adding  $5 + 9$ ). Participants in the shortcut-first condition solved the same problem set but began with the shortcut problems. Consistent with laboratory studies of mental set, participants in the complex-first condition were less likely to use the more efficient shortcut strategy when possible. In addition, these participants were less likely to demonstrate procedural flexibility and conceptual understanding on a subsequent assessment of mathematical equivalence knowledge. These findings suggest that certain problem-solving contexts can help or hinder both flexibility in strategy use and deeper conceptual thinking about the problems.

**Keywords** Mental set · Flexibility · Mathematics · Mathematical equivalence · Conceptual knowledge

✉ Marci S. DeCaro  
marci.decaro@louisville.edu

<sup>1</sup> University of Louisville, Louisville, KY, USA

<sup>2</sup> Department of Psychological and Brain Sciences, 2301 S. Third Street, Louisville, KY 40292, USA

Students often memorize procedures to solve mathematics problems without understanding why these procedures work. Then, students mindlessly apply these procedures to solve new problem types—even when these strategies are no longer efficient or correct (see Baroody, 2003; Langer, 2000; Markovits & Sowder, 1994; McNeil & Alibali, 2005). This common problem is addressed by mathematics curriculum reform standards, which underscore the need for students to understand and flexibly adapt strategies when encountering new problem situations (Kilpatrick, Swafford, & Findell, 2001; National Governors Association, 2010; National Council of Teachers of Mathematics, 2000; Woodward et al., 2012). The goal for students is greater *procedural flexibility*, defined as (a) the knowledge of multiple solution strategies and (b) the use of multiple strategies, particularly in order to select the most efficient strategy to solve a given problem (Rittle-Johnson & Star, 2011; Rittle-Johnson, Star, & Durkin, 2012; Star & Newton, 2009; Verschaffel, Luwelle, Torbeyns, & Van Dooren, 2009). The latter is thought to build upon the former—students sometimes report knowledge of multiple procedures before using these procedures (Blöte, Van der Burg, & Klein, 2001; Durkin, Rittle-Johnson, & Star, 2011; Siegler & Crowley, 1994). By knowing and using multiple solution strategies, students are better able to work with problems beyond a limited range (Siegler, 2003). Students can also reduce the burden on working memory resources by using efficient strategies to solve problems quickly and accurately (Beilock & DeCaro, 2007).

Greater procedural flexibility is also associated with increased understanding of underlying concepts (see Schneider, Rittle-Johnson, & Star, 2011; Verschaffel et al., 2009). When students do not know or use multiple solution strategies, they attend to, or mentally represent, less information about the problem. This limited problem representation thereby constrains how students interpret this information

(McNeil & Alibali, 2005). In contrast, when students are exposed to new problem-solving strategies, they attend to features of the problems that are important for executing these strategies, which helps to support understanding of the underlying concepts (Alibali, Phillips, & Fischer, 2009; McNeil & Alibali, 2004). Numerous studies have demonstrated that knowledge of a variety of solution strategies is associated with deeper conceptual understanding of how these strategies are used (see Baroody, 2003; Heinze, Star, & Verschaffel, 2009; Schneider et al., 2011; Siegler, 1994; Verschaffel et al., 2009).

Although procedural flexibility is associated with more efficient strategy selection and greater conceptual understanding, certain problem-solving contexts can help or hinder such flexibility. For example, elementary-school students often do not notice or use the inversion principle when solving problems of the form  $A + B - B = ?$ ; that is, they fail to recognize that adding or subtracting the same number leaves the original number unchanged (Gilmore & Papadatou-Pastou, 2009; Siegler, 2003). Instead, many students compute each step of these problems, taking longer to perform these calculations and increasing the possibility of error during a given step (Siegler, 2003). However, providing students with a problem-solving context with a large proportion of inversion problems, relative to problems such as  $A + B - C = ?$ , increases their discovery and use of the more efficient inversion-based strategy (Siegler & Stern, 1998; Siegler, 2000; Stern, 1992, 1993).

Numerous laboratory studies on mental set have shown similar effects of the problem context on rigid versus flexible strategy selection. For example, Luchins (1942) provided participants with novel water-jug problems that could all be solved using the same complex strategy (i.e.,  $B - A - 2C$ ). The first set of problems was solvable only by using this complex strategy. However, subsequent problems could be solved either by using this same strategy or by using a much simpler strategy (e.g.,  $A + C$ ). After solving several problems using the complex strategy, individuals generally continued using this strategy, even when the simpler procedure became available to use. The early experience of solving problems using the complex strategies led individuals to select less efficient strategies, the hallmark of mental set (see also Beilock & DeCaro, 2007; Wiley, 1988).

## Current studies

Thus, separate lines of research have demonstrated that (a) the problem-solving context can influence procedural flexibility and mental set, and (b) procedural flexibility is related to conceptual understanding. The current studies examined the consequences of inducing mental set on both procedural flexibility and activation of the underlying concepts. Specifically, mental set was manipulated by providing participants with

mathematics problems that could only be solved using a complex strategy. It was expected that participants would persist in using this strategy even when, later, the problem features changed, better allowing for more efficient strategies to be used. It was also predicted that such procedural inflexibility would impact individuals' thinking about the problems. Mental set occurs because individuals narrowly represent the problems—overlooking features that may indicate the need to adapt one's strategy (e.g., McNeil & Alibali, 2005; McNeil, Rittle-Johnson, Hattikudur, & Petersen, 2010; Wiley, 1988). Thus, when mental set is induced, participants may also demonstrate less understanding of the concepts underlying the problems. Rigid thinking may therefore be revealed in measures of both problem-solving flexibility and conceptual understanding.

These predictions were examined with undergraduate students in a laboratory setting (Experiment 1) and with children in a classroom setting (Experiment 2) in the domain of mathematical equivalence. This domain was selected because there are well-established measures that can be used to determine multiple types of understanding (e.g., procedural knowledge, conceptual knowledge). In addition, mathematical equivalence problems (e.g.,  $5 + 2 + 3 = \_ + 3$ ) can be solved using multiple possible methods, some of which are more efficient than others. One strategy students commonly use is known as the *add-subtract* strategy, in which students add up the values on the left side of the equal sign and subtract the value(s) on the right side. Students can also use shortcut strategies that can be applied to many, though not all, problems. For example, the *grouping* strategy may be used when both sides of the problem have a repeated addend (e.g., the "3" in the problem  $5 + 2 + 3 = \_ + 3$ ). Use of this procedure requires eliminating the repeated addend and solving the problem without these values (e.g.,  $5 + 2 = \_$ ). By using the grouping procedure, the problem can be solved with fewer operations. Thus, fewer numbers must be manipulated to compute the answer, reducing the cognitive demand of the task and enabling the problem to be solved more quickly and accurately. For these reasons, grouping can be considered a more efficient problem-solving strategy (cf. Rittle-Johnson et al., 2012). Flexibly using the shortcut strategy when it is available might not only increase procedural efficiency but might also indicate understanding of the underlying concept of mathematical equivalence—that the equal sign represents a relational symbol.

The current study tested the idea that the context in which mathematical equivalence problems are solved can influence whether students demonstrate problem-solving flexibility and conceptual understanding. Participants solved mathematical equivalence problems in one of two conditions. In both conditions, participants were given two sets of problems to solve, labeled here as complex problems and shortcut problems. *Complex problems* (e.g.,  $7 + 5 + 9 = 3 + \_$ ) did not include a repeated addend on both

sides of the equal sign and therefore were expected to be solved using a complex procedure (e.g., add–subtract). *Shortcut problems* (e.g.,  $7 + 5 + 9 = 7 + \_$ ) had a repeated addend and therefore could be solved using a shortcut strategy (i.e., grouping). The two conditions differed only in the order in which these problem sets were presented. Participants in the *complex-first condition* solved complex problems followed by shortcut problems, and participants in the *shortcut-first condition* solved the problem sets in the reverse order. It was hypothesized that the complex-first condition would induce mental set—reducing use of the shortcut strategy. It was also predicted that this mental set would reduce performance on subsequent measures of procedural flexibility and conceptual understanding.

Such findings would indicate that anchoring students on complex solution methods may not only limit their ability to notice and use more efficient strategies when possible but may also prompt them to become more rigid in their thinking about the underlying concepts. These findings may further shed light on one potential factor reducing students' flexibility and understanding in the mathematics classroom—traditional methods of instruction that anchor students on complex procedures before introducing shortcut strategies.

## Experiment 1

### Method

#### Participants

Participants ( $N = 41$ , 83% female) were undergraduate students from the psychology subject pool who received course credit or payment. Two participants were excluded from the data set for scoring below 85% correct on the set problems, indicating they did not attain mental set (see Beilock & DeCaro, 2007; Gasper, 2003; Schultz & Searleman, 1998). Five additional participants were excluded because they reported that they were bilingual, which has been demonstrated to improve flexibility on problem-solving tasks (Cushen & Wiley, 2011).

#### Procedure

Participants were tested in groups of one to five. Following informed consent, each participant was given a packet with the set problems, followed by the mathematical equivalence assessment. The experimenter instructed participants not to go backwards in the packet. Finally, participants completed demographic questions and were debriefed. The session lasted approximately 20 to 30 minutes.

### Set problems

Set problems were mathematical equivalence problems (see Table 1). Participants were instructed to find the number that goes in each box to make the amount on the left side of the equal sign the same as the amount on the right side of the equal sign. They were asked to show their work and to try to use fast and correct ways to solve the problems.

The set problems were divided into two sections, and the order in which these sections were presented differed depending on experimental condition. Participants in the complex-first condition ( $n = 20$ ) were first given six complex problems without a repeated addend, for which only complex (i.e., multistep) strategies could be used (see Table 1). Following, these participants solved six shortcut problems, for which one addend was repeated on both the left and right sides of the equal sign (see Table 1). Thus, these problems could be solved using either a complex strategy or a shortcut strategy requiring fewer steps (i.e., the grouping strategy). The complex and shortcut problems included the same numbers, except for the presence of the repeated addend, to control for difficulty across the two problem sets. Participants in the shortcut-first condition ( $n = 21$ ) completed the exact same problem sets but in the reverse order—six shortcut problems followed by six complex problems. Thus, the only difference between conditions was whether the shortcut problems were solved first or last. Of key interest was whether participants used the more efficient shortcut strategy to solve problems for which the shortcut was available, and whether condition impacted performance on the mathematical equivalence assessment.

#### Mathematical equivalence assessment

The mathematical equivalence assessment was adapted from previous research (DeCaro & Rittle-Johnson, 2012; Matthews, Rittle-Johnson, McEldoon, & Taylor, 2012; Rittle-Johnson, Taylor, Matthews, & McEldoon, 2011). The assessment included four subscales (see Table 2): *Conceptual Knowledge* items evaluated both individuals' understanding of the meaning of the equal sign as a relational symbol and the structure of equations. *Procedural Knowledge* items assessed participants' ability to solve mathematical equivalence problems. The problems were of the same structure as the set items, and problems with and

**Table 1** Set problems used in Studies 1 and 2

Complex problems	Shortcut problems
$7 + 5 + 9 = 3 + \square$	$7 + 5 + 9 = 7 + \square$
$4 + 14 + 8 = 6 + \square$	$4 + 14 + 8 = 4 + \square$
$15 + 3 + 9 = 13 + \square$	$15 + 3 + 9 = 15 + \square$
$9 + 7 + 6 = \square + 3$	$9 + 7 + 6 = \square + 6$
$14 + 5 + 3 = \square + 7$	$14 + 5 + 3 = \square + 3$
$6 + 3 + 12 = \square + 15$	$6 + 3 + 12 = \square + 12$

**Table 2** Example mathematical equivalence assessment items

Item	Scoring criteria
Conceptual Knowledge (4 items in Exp. 1, $\alpha = .24^a$ ; 6 items in Exp. 2, $\alpha = .62$ )	
What does the equal sign (=) mean?	1 point if defined relationally (e.g., “both sides are the same”)
Rate definitions of the equal sign as “good,” “not good,” or “don’t know”	1 point if rated the statement “The equal sign means the same as” as a good definition.
Procedural Knowledge (9 items; $\alpha = .41$ in Exp. 2 <sup>b</sup> )	
Find the number that goes in each box. . . . Try to use fast (and correct) ways to solve the problems. (e.g., $7 + 14 + 5 = 7 + \square$ ; $11 + 7 + 8 = \square + 13$ )	1 point for correct answer ( $\pm 1$ )
Transfer (3 items; $\alpha = .64$ in Exp. 1; $\alpha = .42$ in Exp. 2)	
Without subtracting the 7, can you tell if the number sentence is true or false? $56 + 85 = 141$ is true. Is $56 + 85 - 7 = 141 = 7$ true or false? How do you know?	1 point if answer demonstrates knowledge that performing the same task on each side of the equal sign maintains equivalence
Find the number that goes in each box. You can try to find a shortcut so you don’t have to do all the adding. Show your work and write your answer in the box. $898 + 13 = 896 + \square$ ; $43 + \square = 48 + 76$	1 point for correct answer ( $\pm 1$ )
Flexibility Knowledge (4 items; $\alpha = .86$ in Exp. 1; $\alpha = .82$ in Exp. 2)	
Find the number that goes in each box, using two DIFFERENT ways of getting the answer. Way 1 Way 2 $5 + 9 + 11 = 5 + \square$ $5 + 9 + 11 = 5 + \square$	1 point for two correct, unique solutions

<sup>a</sup> This low Cronbach’s alpha was driven by the one item assessing knowledge of the structure of equations, which was slightly negatively correlated with the other three items in the conceptual knowledge subscale ( $r$ s ranged from  $-.094$  to  $-.036$ ). Excluding this item from the scale raises alpha to a more moderate value of  $.41$ . This item was included in the conceptual knowledge subscale to be consistent with previous research (e.g., DeCaro & Rittle-Johnson, 2012) and with Experiment 2. In addition, excluding this item did not change the pattern of results for the conceptual knowledge subscale, which remained statistically significant.

<sup>b</sup> Cronbach’s alpha could not be computed for the procedural knowledge subscale in Experiment 1, because six of the nine items had 100% accuracy and thereby zero variance.

without repeated-addends (seven items and two items, respectively) were intermixed. *Transfer* was assessed with items requiring a higher level of conceptual thinking. Procedural flexibility was assessed in two ways. *Flexible Knowledge* of solution strategies was assessed using items adapted from research on algebra equation solving (Rittle-Johnson et al., 2012), requiring generation and evaluation of different strategies to solve problems. *Flexible use* of procedures was assessed by coding the strategies used to solve correct Procedural Knowledge items and calculating the number of different correct strategies used (see Table 3).

### Coding

Items requiring a written explanation were coded, and interrater reliability was assessed by a second coder for 20% of the participants. Interrater agreement was high (percent agreement = 97%–100%). In addition, correct strategies used to solve the set problems and the Procedural Knowledge items were inferred from participants’ written work and coded using the categories presented in Table 3. Strategies for 20% of the participants were coded by two raters, and interrater agreement was high (percent agreement = 93%–96%).

**Table 3** Strategy codes for correct mathematical equivalence problems

	Description	Example
		$4 + 5 + 8 = \square + 8$
<i>Add-subtract</i>	Adds the numbers on one side of the equation and then subtracts the number on the other side	$4 + 5 + 8 = 17$ $17 - 8 = 9$
<i>Equalizer</i>	Adds the numbers on one side of the equation and determines what number goes in the box to make the other side the same; has a clear understanding that two sides must be equal	$4 + 5 + 8 = 17$ $9 + 8 = 17$
<i>Grouping</i>	Ignores the repeated addend and adds the other numbers	$4 + 5 = 9$
<i>Insufficient work</i>	The answer is correct. Either no work or partial work is shown, or the work that is shown is ambiguous.	

## Results and discussion

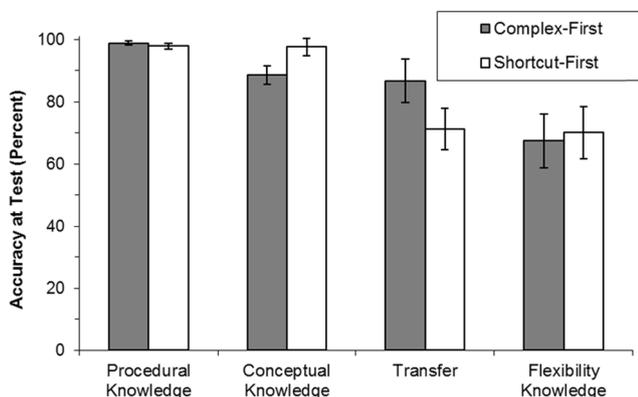
### Set problems

Accuracy on the set problems was high in both the shortcut-first ( $M = 97.70\%$ ,  $SD = 0.70\%$ ) and complex-first ( $M = 98.80\%$ ,  $SD = 0.70\%$ ) conditions,  $F(1, 39) = 1.08$ ,  $p = .306$ . Strategy use, however, was significantly different between conditions. Compared to participants in the shortcut-first condition (56%), participants in the complex-first condition (19%) were less likely to use the grouping shortcut strategy to solve the set problems,  $\chi^2(1, N = 34) = 4.86$ ,  $p = .028$ . Data from seven participants (complex-first condition,  $n = 4$ ; shortcut-first condition,  $n = 3$ ) were excluded from this analysis because 80% or more of their strategies could not be coded (i.e., fell in the “insufficient work” category; see Table 3). The findings for the set problems establish that the complex-first condition increased mental set relative to the shortcut-first condition.

### Mathematical equivalence assessment

Next, analyses were conducted to examine whether inducing mental set impacted performance on a subsequent test of mathematical equivalence knowledge. The percentage of accuracy on the four mathematical equivalence assessment subscales was examined as a function of condition by conducting separate ANOVAs. As shown in Fig. 1, no significant difference between conditions was found for Procedural Knowledge,  $F < 1$ , Transfer,  $F(1, 39) = 2.50$ ,  $p = .122$ , or Flexibility Knowledge,  $F < 1$ . However, participants in the complex-first condition scored at a significantly lower level than those in the shortcut-first condition on the Conceptual Knowledge subscale,  $F(1, 39) = 4.67$ ,  $p = .037$ ,  $\eta_p^2 = .11$ .

In addition to the Flexibility Knowledge subscale, we examined the flexible use of procedures, which is a more advanced assessment of flexibility thought to build on flexibility knowledge (Rittle-Johnson et al., 2012; Star & Newton, 2009).



**Fig. 1** Percent correct on the four mathematical equivalence test subscales as a function of set condition in Experiment 1. Error bars represent standard errors

Flexible use of procedures was calculated as the number of different correct strategies used on the Procedural Knowledge items (see Table 3). Data from six participants were cut from this analysis for having 80% or more strategies that could not be coded (complex-first condition,  $n = 2$ ; shortcut-first condition,  $n = 4$ ). As shown in Fig. 2, participants in the complex-first condition used significantly fewer correct strategies at posttest than participants in the shortcut-first condition,  $F(1, 33) = 4.91$ ,  $p = .034$ ,  $\eta_p^2 = .13$ , indicating less flexible use of procedures. As shown in Table 4, this effect appears to be driven by a decreased use of the grouping strategy in the shortcut-first condition ( $M = 39.87$ ,  $SD = 36.65$ ; CI [21.02, 58.71]); complex-first condition ( $M = 20.37$ ,  $SD = 32.84$ ),  $d = .56$ .

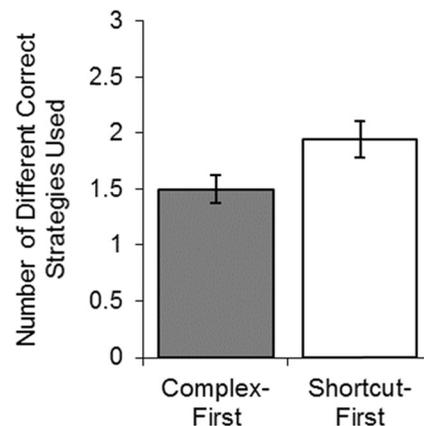
Together, these findings support the hypothesis that mental set can be induced in the context of mathematical equivalence problems. More importantly, these findings demonstrate that such mental set limits conceptual understanding and subsequent flexibility in strategy use. Experiment 2 aimed to replicate and extend these results to a classroom setting, with elementary-school children. In addition, a measure of working memory capacity was administered. Working memory capacity predicts performance on an array of academic skills, including mathematical problem solving (e.g., Barrett, Tugade, & Engle, 2004; Bull, Espy, & Wiebe, 2008; Swanson, 2011). This measure was included to determine that the impact of set condition occurs regardless of cognitive ability measures (i.e., for both low- and high-capacity students).

## Experiment 2

### Method

#### Participants

Participants ( $N = 87$ , 48% female) were fifth-grade ( $n = 64$ ) and sixth-grade ( $n = 23$ ) students from two private schools



**Fig. 2** Flexible use of procedures on the mathematical equivalence test in Experiment 1 as a function of condition. Error bars represent standard errors

**Table 4** Mean percent correct strategy use on posttest problems by experiment and condition (95% CIs in brackets)

	Experiment 1		Experiment 2	
	<i>Complex-first condition</i>	<i>Shortcut-first condition</i>	<i>Complex-first condition</i>	<i>Shortcut-first condition</i>
<i>Grouping</i>	20.37 [4.04, 36.70]	39.87 [21.02, 58.71]	11.97 [3.05, 20.88]	23.04 [11.80, 34.27]
<i>Add-subtract</i>	40.12 [17.85, 62.40]	31.37 [14.57, 48.18]	53.28 [38.28, 68.27]	55.28 [41.14, 69.43]
<i>Equalizer</i>	27.78 [10.08, 45.48]	20.26 [3.82, 36.70]	25.36 [11.78, 38.93]	19.24 [7.35, 31.14]

in Kentucky and Tennessee (fifth-grade only) for whom parental consent and student assent were obtained. Additional students were tested but not included in the study for scoring below 85% correct on the set problems ( $n = 22$ ). This sample was chosen to ensure that students in the dataset were able to solve the problems and thereby attain mental set (see Beilock & DeCaro, 2007; Gasper, 2003; Schultz & Searleman, 1998). Three additional participants were excluded because they did not finish the mathematical equivalence test packet.

#### *Procedure*

Participants completed the study in their mathematics classes. An experimenter read all instructions, and students worked individually. Students completed the set problems, followed by the mathematical equivalence assessment. Students were instructed not to go backwards in their packet. The session lasted approximately 30 to 45 minutes. On a separate day within a week of the first session, students were given the working memory assessment individually in a quiet location in their school.

#### *Set problems*

Set problems were the same mathematical equivalence problems given in Experiment 1 (see Table 1). As in Experiment 1, students were asked to show their work and to try to use fast and correct ways to solve the problems. Participants were randomly assigned to work on different packets within the same classroom, depending on condition. Students in the complex-first condition ( $n = 43$ ) completed the complex problems first, followed by the shortcut problems. Students in the shortcut-first condition ( $n = 42$ ) completed the shortcut problems, followed by the complex problems.

#### *Mathematical equivalence assessment*

The mathematical equivalence assessment was the same as that used in Experiment 1, with one exception. Two additional Conceptual Knowledge items were included, based on previous research (e.g., DeCaro & Rittle-Johnson, 2012; Rittle-

Johnson & Alibali, 1999). For these two items, students were shown a mathematical equivalence problem for 5 seconds and were asked to write down the problem from memory. These items assess students' encoding of the problem structure, such as the presence of numbers on both sides of the equal sign (McNeil & Alibali, 2004). Also in contrast to Experiment 1, each section was given an approximate time deadline that was not announced to students. However, the experimenter did not move on in the packet until the majority of the students were finished with a section. Students were instructed to wait at the end of each section and to move forward on the assessment as a class. Then, the instructions for each section were read aloud to students.

#### *Working memory measure*

Working memory capacity was measured using the backwards digit span task from the *Wechsler Intelligence Scale for Children* (WISC-IV) Working Memory Index (Wechsler, 2003). In this task, individual students are read a series of numbers at a rate of one per second and asked to repeat the numbers in reverse order. Number series lengths begin at two and end at a maximum of eight. There are two items per series length. The task is discontinued when a student recalls both items in a series of a given length incorrectly. Working memory scores consist of the number of series that the student correctly recalls in backward order. The forward digit span task, a measure of short-term memory, was administered prior to the backwards digit span, to provide students with practice on the procedure and to be consistent with prior methods of administration. However, scores on the forward digit span task were not relevant to this study and were not examined further.

#### *Coding*

As in Experiment 1, two raters independently coded 20% of items requiring a written explanation, and the percentage of agreement was high across subscales (90%–95%). In addition, written strategies were coded by two raters, using the categories presented in Table 3. Percentage of agreement was high (95%–98%).

## Results and discussion

### Working memory capacity

Two participants did not complete the backwards digit span, which removed their data from all analyses including this variable. Working memory scores did not differ between conditions (complex-first condition:  $M = 5.56$ ,  $SD = 1.79$ ; shortcut-first condition:  $M = 5.19$ ,  $SD = 1.70$ ),  $F < 1$ , indicating that random assignment was successful. In addition, preliminary regression analyses including working memory (centered), condition (effects coded), and a Working Memory  $\times$  Condition interaction term demonstrated that working memory capacity did not interact with condition to predict any of the primary dependent measures ( $\beta$ s =  $-.125$  to  $.103$ ,  $p$ s =  $.379$  to  $.657$ ). Collapsing across condition, higher working memory capacity was associated with better performance on two mathematical equivalence subscales: Conceptual Knowledge,  $r(85) = .291$ ,  $p = .007$ , and Transfer,  $r(85) = .310$ ,  $p = .004$ . Working memory capacity was not significantly correlated with accuracy on the other subscales: Procedural Knowledge,  $r(85) = .198$ ,  $p = .070$ ; Flexibility Knowledge,  $r(85) = .152$ ,  $p = .164$ ; flexible use of strategies,  $r(79) = .179$ ,  $p = .114$ . Working memory capacity was also not correlated with use of the grouping strategy on the set problems,  $r(82) = .130$ ,  $p = .244$ .

Thus, working memory capacity was associated with better performance on Conceptual Knowledge and Transfer, regardless of condition. Working memory capacity was included as a covariate in the analyses of condition reported below, to ensure that effects of condition occur beyond any influence of working memory ability. Removing this covariate from analyses does not change the pattern of results.

### Set problems

Students performed at an equally high level on the set problems across both conditions (complex-first condition:  $M = 96.51\%$ ,  $SD = 4.16\%$ ; shortcut-first condition:  $M = 97.02\%$ ,  $SD = 4.04\%$ ),  $F < 1$ . However, students in the complex-first condition were less likely to use the grouping shortcut strategy to solve the set problems,  $\chi^2(1, N = 66) = 5.21$ ,  $p = .022$ . Ten percent of students in the complex-first condition used the shortcut strategy, compared to 30% of students in the shortcut-first condition. Data from four students (complex-first condition,  $n = 3$ ; shortcut-first condition,  $n = 1$ ) were excluded from this analysis because 80% or more of their strategies could not be coded.

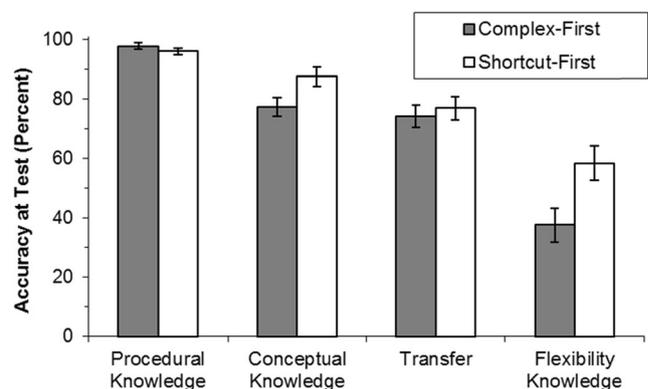
Thus, as in Experiment 1, students demonstrated greater mental set in the complex-first condition. Moreover, these findings demonstrate that mental set can be instantiated based

on a single mathematical problem-solving experience, with children, in an educational setting.

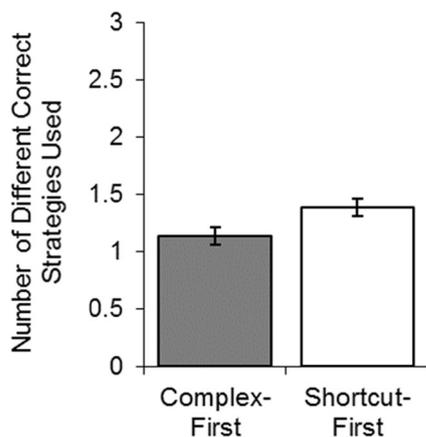
### Mathematical equivalence assessment

As shown in Fig. 3, no difference between conditions was found for two of the math equivalence subscales: Procedural Knowledge,  $F(1, 82) = 1.11$ ,  $p = .295$ , and Transfer,  $F < 1$ . In contrast, a significant effect of condition was found for two subscales: Conceptual Knowledge,  $F(1, 82) = 5.13$ ,  $p = .026$ ,  $\eta_p^2 = .06$ , and Flexibility Knowledge,  $F(1, 82) = 6.54$ ,  $p = .012$ ,  $\eta_p^2 = .07$ . Students in the complex-first condition scored at a lower level than students in the shortcut-first condition. In addition, as shown in Fig. 4, students in the complex-first condition demonstrated less flexible use of procedures on the test problems,  $F(1, 76) = 5.29$ ,  $p = .024$ ,  $\eta_p^2 = .065$ . As shown in Table 4, this effect was driven by a decreased use of the grouping strategy in the complex-first condition ( $M = 11.97$ ,  $SD = 27.50$ ; CI [3.05, 20.88]); shortcut-first condition ( $M = 23.04$ ,  $SD = 35.61$ ),  $d = .35$ .

These results replicate those of Experiment 1, demonstrating that instantiating mental set leads to worse performance on subsequent measures of conceptual understanding and less flexible use of strategies. Students in Experiment 2 showed an additional detriment in the complex-first condition for flexibility knowledge. Flexible use of strategies is thought to build on flexibility knowledge (Rittle-Johnson et al., 2012; Star & Newton, 2009). Compared to the undergraduate student sample in Experiment 1, students in Experiment 2 likely had less experience solving mathematical equivalence problems. It is not surprising, then, that the younger students would demonstrate less flexibility knowledge, particularly when primed by complex strategies. Such findings indicate that mental set may be even more detrimental to procedural flexibility for younger students with less experience with the mathematics domain, by limiting both knowledge and flexible use of multiple procedures.



**Fig. 3** Percent correct on the four mathematical equivalence test subscales as a function of set condition in Experiment 2. Error bars represent standard errors



**Fig. 4** Flexible use of procedures on the mathematical equivalence test in Experiment 2 as a function of condition. Error bars represent standard errors

## General discussion

In two experiments, mental set was induced in the context of mathematics problems. Undergraduate (Experiment 1) and fifth- and sixth-grade (Experiment 2) student participants solved a set of mathematical equivalence problems in one of two conditions. In the complex-first condition, participants first solved problems without a repeated addend (e.g.,  $7 + 5 + 9 = 3 + \_$ ) that required multiple steps to complete (e.g., using the add-subtract strategy). Then, these participants solved problems with a repeated addend (e.g.,  $7 + 5 + 9 = 7 + \_$ ), for which a simpler shortcut strategy could be used (e.g., the grouping strategy). Participants in the shortcut-first condition solved these same problem sets but began with the shortcut problems. In both experiments, participants in the complex-first condition were more likely to use a complex strategy to solve the mathematical equivalence problems when a shortcut was available. These findings mirror classic laboratory mental set effects (e.g., Luchins, 1942; Wiley, 1988) using an ecologically valid mathematics task.

These findings also extend prior research to examine the impact of mental set on subsequent assessments of mathematics knowledge and procedure use. In both experiments, no condition differences were found on measures of problem-solving ability or transfer. However, participants for whom mental set was induced (complex-first condition) demonstrated less conceptual understanding and procedural flexibility, in that they were less likely to use multiple procedures to solve mathematical equivalence problems given on a separate post-test. In Experiment 2, participants also demonstrated less knowledge of multiple correct solution strategies in the complex-first condition. Interestingly, except for this difference, undergraduates (Experiment 1) showed the same pattern of findings as children (Experiment 2). Other work has similarly shown that undergraduates sometimes use less advanced strategies than might be otherwise expected (see McNeil &

Alibali, 2005; McNeil et al., 2010). Here, we see that similar to children, undergraduates can fall prey to the negative effects of a mathematics problem-solving context, even if they otherwise know the domain well.

These findings extend laboratory research on mental set, demonstrating that the effects of set can reach beyond problem-solving strategies to one's conceptual understanding of the problem domain. These findings suggest that mental set has implications for the ways in which the problems are conceptually represented (e.g., Alibali et al., 2009). Individuals who become used to solving problems a certain way may overlook important problem features (e.g., the presence of a repeated addend on both sides of the equal sign; Alibali et al., 2009; McNeil & Alibali, 2004, 2005), which may limit both flexible use of strategies and consideration of the underlying mathematics concepts (e.g., that the equal sign means that two quantities are the same). This conceptual link between procedural flexibility and conceptual understanding is consistent with a great deal of correlational evidence (see Schneider et al., 2011). The current work demonstrates a causal impact of inducing mental set on both procedural flexibility and conceptual understanding.

## Working memory capacity

It is of interest that working memory capacity did not moderate the effects of condition (Experiment 2). Higher working memory capacity is associated with better performance on a range of skills, including academic tasks such as mathematical problem solving (e.g., Barrett et al., 2004; Bull et al., 2008; Swanson, 2011). This effect was generally present in the current results as well. Higher working memory predicted higher scores on the Conceptual Knowledge and Transfer subscales and was positively, but not significantly, associated with scores on the remaining subscales. However, working memory capacity did not impact susceptibility to mental set in this study. Indeed, the effects of condition remained when controlling for working memory capacity, suggesting that the effects of mental set occurred regardless of individual differences in this cognitive ability.

The sample used in Experiment 2 may have limited the ability to find effects of working memory capacity on procedural flexibility. Specifically, working memory capacity was measured with children who solved the set problems at a relatively high rate of accuracy. One possible prediction would be that individuals with lower working memory capacity may be less capable of switching strategies in the complex-first condition, due to reduced ability to flexibly switch task sets (cf. Bull, Johnston, & Roy, 1999; Bull & Scerif, 2001). However, such findings are typically seen with children with lower mathematics ability (e.g., Bull & Scerif, 2001), who may have been largely excluded from the current sample. An opposite prediction could be that individuals with higher

working memory capacity would be less capable of switching strategies in the complex-first condition, because of a tendency to persist in using complex strategies (e.g., Beilock & DeCaro, 2007; DeCaro, Thomas, & Beilock, 2008). However, such findings have been shown with adults, who demonstrate generally higher working memory capacity than children (cf. Alloway, 2006), and may therefore be more likely to utilize these resources. More work is needed to determine when working memory capacity will either enable or reduce flexibility and subsequent performance outcomes.

### Strategy efficiency

Mental set in this study led to less demonstrated conceptual knowledge and, in Experiment 2, less demonstrated knowledge of multiple procedures—both inarguably less than optimal performance outcomes. Mental set also resulted in less flexible use of procedures to solve the problems. However, this pattern of behavior may not necessarily be inefficient. Although the shortcut (grouping) strategy in this study required fewer computations, problem-solving accuracy across all problems was very high. Thus, although unmeasured, the primary benefits to using the shortcut were presumably in reduced time and cognitive demand to solve the problems.

Yet, there is also a cost to switching strategies. Specifically, research has demonstrated that switching between complex and shortcut strategies incurs a cost to reaction time (e.g., Lemaire & Lecacheur, 2010; Luwel, Schillemans, Onghena, & Verschaffel, 2009; Schillemans, Luwel, Bulté, Onghena, & Verschaffel, 2009). This cost likely reflects both interference from a previously used strategy and a need to use executive control processes to activate the new strategy and deactivate the previously used strategy (Lemaire & Lecacheur, 2010). Thus, it is possibly equally efficient to continue using a complex strategy, even when a shortcut is available. More research is needed, to determine the relative efficiency of switching to shortcut strategies on outcomes such as accuracy, reaction time, and cognitive load.

What is less debatable, however, is the finding that a failure to use shortcut strategies on set problems limits one's thinking about the problems. Thus, even though flexible use of strategies might not always be inefficient, when considering the costs associated with switching, such inflexibility may bias one's conceptual understanding in addition to one's problem-solving approaches.

### Implications for educational practice

These findings also have potential implications for educational practice, suggesting that the context in which students solve problems may impact their understanding of the problems. Strategy selection is often automatic, rather than executed via deliberative choice (Siegler & Jenkins, 1989; Verschaffel

et al., 2009). Students in our experiments had prior knowledge of solution strategies, yet priming certain strategies impacted their strategy use in one experimental session. In educational practice, students are often taught complex strategies exclusively, and only later (sometimes) are taught more efficient solution methods (Hiebert et al., 2003). As students gain more experience with a particular strategy, they are increasingly less likely to use other strategies (McNeil & Alibali, 2005; Schillemans et al., 2009; Siegler, 2002). Thus, early, focused instruction on complex strategies may unnecessarily lead to mental set over time, or an emphasis on complex solution methods. Consequently, students may overlook important problem features that enable them to flexibly adapt these procedures (e.g., to use conceptually driven shortcuts) and understand the underlying mathematical concepts (Verschaffel et al., 2009). Future research is needed to investigate the impact of teaching students new mathematics strategies on the establishment of mental set and conceptual understanding over time.

One area of research within the mathematical equivalence domain is consistent with these ideas. Previous studies have demonstrated that students can become “entrenched” in using arithmetic procedures that lead them to overlook important features of mathematical equivalence problems (namely, the presence of addends on both sides of the equal sign; Alibali et al., 2009; McNeil & Alibali, 2004). Specifically, children often begin solving mathematics problems with the equal sign at the end of the problem (e.g.,  $7 + 3 = \_$ ). These early experiences often lead students to misperceive the equal sign as meaning “get the answer” or “find the total” (Baroody & Ginsburg, 1983; Carpenter, Franke, & Levi, 2003; McNeil & Alibali, 2005). This misperception leads to difficulty in learning to solve and understand mathematical equivalence problems. For example, students will give the answer “10” to a problem such as “ $7 + 3 = \_ + 4$ .” Thus, early experience with certain problem-solving contexts can lead students to apply incorrect strategies to solve problems that differ in important features (McNeil, Fyfe, & Dunwiddie, 2015; McNeil et al., 2010). The current studies go beyond this work to suggest that certain problem-solving experiences may also limit flexibility in applying multiple correct problem-solving procedures, as well as conceptual understanding, even after students demonstrate an ability to solve problems in this domain. Thus, the current findings offer potential new insights for educational practice (see Mayer, 2001), while providing new information about the impact of problem context on conceptual representation.

**Author Notes** The author thanks Bethany Rittle-Johnson, Katherine McEldoon, Joanna Weaver, and Natalie Hummel for helpful discussion and assistance with this research as well as the staff, teachers, and children at University School of Nashville, TN, and St. Nicholas Academy of Louisville, KY.

## References

- Alibali, M. W., Phillips, K. M. O., & Fischer, A. D. (2009). Learning new problem solving strategies leads to changes in problem representation. *Cognitive Development, 24*, 89–101. doi:10.1016/j.cogdev.2008.12.005
- Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews, 1*, 134–139.
- Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 1–33). Mahwah, NJ: Erlbaum.
- Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal, 84*, 199–212.
- Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553–573. doi:10.1037/0033-2909.130.4.553
- Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 983–998.
- Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of Educational Psychology, 93*, 627–638.
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205–228. doi:10.1080/87565640801982312
- Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*, 421–442.
- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19*, 273–293.
- Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in the elementary school*. Portsmouth, NH: Heinemann.
- Cushen, P. J., & Wiley, J. (2011). Aha! Voila! Eureka! Bilingualism and insightful problem solving. *Learning and Individual Differences, 21*, 458–462. doi:10.1016/j.lindif.2011.02.007
- DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology, 113*, 552–568. doi:10.1016/j.jecp.2012.06.009
- DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284–294.
- Durkin, K., Rittle-Johnson, B., & Star, J. R. (2011, August). *Procedural flexibility matters for student achievement: How procedural flexibility relates to other outcomes*. Paper presented at the biennial meeting of the European Association for Research on Learning and Instruction, Exeter, England.
- Gasper, K. (2003). When necessity is the mother of invention: Mood and problem solving. *Journal of Experimental Social Psychology, 39*, 248–262.
- Gilmore, C. K., & Papadatou-Pastou, M. (2009). Patterns of individual differences in conceptual understanding and arithmetical skill: A meta-analysis. *Mathematical Thinking and Learning, 11*, 25–40. doi:10.1080/10986060802583923
- Heinze, A., Star, J. R., & Verschaffel, L. (2009). Flexible and adaptive use of strategies and representations in mathematics education. *ZDM Mathematics Education, 41*, 535–540. doi:10.1007/s11858-009-0214-4
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., ... Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study (NCES, No. 2003-013)*. Washington, DC: US Department of Education, National Center for Education Statistics.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
- Langer, E. J. (2000). Mindful learning. *Current Directions in Psychological Science, 9*, 220–223.
- Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition, 38*, 322–332. doi:10.3758/MC.38.3.322.
- Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs, 54*, 1–95. doi:10.1037/h0093502
- Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology, 100*, 753–771. doi:10.1348/000712609X402801.
- Markovits, Z., & Sowder, J. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education, 25*, 4–29.
- Matthews, P. G., Rittle-Johnson, B., McEldoon, K. L., & Taylor, R. T. (2012). Measure for measure: What combining diverse measures reveals about children's understanding of the equal sign as an indicator of mathematical equality. *Journal for Research in Mathematics Education, 43*, 316–350.
- Mayer, R. E. (2001). What good is educational psychology? The case of cognition and instruction. *Educational Psychologist, 36*, 83–88.
- McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451–466. doi:10.1016/j.cogsci.2003.11.002
- McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883–899.
- McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology, 107*, 423–436. doi:10.1037/a0037687
- McNeil, N. M., Rittle-Johnson, B., Hattikudur, S., & Petersen, L. A. (2010). Continuity in representation between children and adults: Arithmetic knowledge hinders undergraduates' algebraic problem solving. *Journal of Cognitive Development, 11*, 437–457. doi:10.1080/15248372.2010.516421
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, DC: Authors.
- Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175–189. doi:10.1037/0022-0663.91.1.175
- Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: Learning outcomes supported by different types of comparisons. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 199–222). Waltham, MA: Elsevier. doi:10.1016/B978-0-12-387691-1.00007-7
- Rittle-Johnson, B., Star, J., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing

- procedures? *British Journal of Educational Psychology*, 82, 436–455. doi:10.1111/j.2044-8279.2011.02037.x
- Rittle-Johnson, B., Taylor, R., Matthews, P. G., & McEldoon, K. (2011). Assessing knowledge of mathematical equivalence. A construct modeling approach. *Journal of Educational Psychology*, 103, 85–104.
- Schillemans, V., Luwel, K., Bulte, I., Onghena, P., & Verschaffel, L. (2009). Influence of previous strategy use on individuals' subsequent strategy choice: Findings from a numerosity judgement task. *Psychologica Belgica*, 49, 191–205.
- Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, 47, 1525–1538. doi:10.1037/a0024997
- Schultz, P. W., & Searleman, A. (1998). Personal need for structure, the Einstellung task, and the effects of stress. *Personality and Individual Differences*, 24, 305–310.
- Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, 3, 1–5.
- Siegler, R. S. (2000). Unconscious insights. *Current Directions in Psychological Science*, 9, 79–83.
- Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Gamott & J. Parziale (Eds.), *Microdevelopment: A process oriented perspective for studying development and learning* (pp. 31–58). Cambridge, MA: Cambridge University Press.
- Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. In J. Kilpatrick, W. B. Martin, & D. E. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 219–233). Reston, VA: National Council of Teachers of Mathematics.
- Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology*, 27, 194–226.
- Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, 127, 377–397.
- Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM—The International Journal on Mathematics Education*, 41, 557–567. doi:10.1007/s11858-009-0185-5
- Stern, E. (1992). Spontaneous use of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology*, 17, 266–277.
- Stern, E. (1993). What makes certain arithmetic word problems involving comparison of sets so difficult for children? *Journal of Educational Psychology*, 85, 7–23.
- Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, 103, 821–837. doi:10.1037/a0025114
- Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, 24, 335–359. doi:10.1007/BF03174765
- Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio, TX: Harcourt Assessment.
- Wiley, J. (1988). Expertise as a mental set. *Memory & Cognition*, 26, 716–730.
- Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., ... & Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide* (NCEE 2012-4055). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.