Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Some Correct Strategies Are Better Than Others: Individual Differences in Strategy Evaluations Are Related to S

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Menendez, David	University of Wisconsin-Madison	
Brown, Sarah	Boston University	
Alibali, Martha	University of Wisconsin-Madison	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2023

Check type of content being submitted and complete one of the following in the box below:

- ☑ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

Cognitive Science, 47, 3

DOI or URL to published work (if available) 10.1111/cogs.13269

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** Institute of Education Scie through **[Grant number]** R305B1500 to **Institution]** University of W. The opinions expressed are those of the authors and do not represent views of the **[Office name]** Institute of Education Scie or the U.S. Department of Education.

Some correct strategies are better than others: Individual differences in strategy evaluations are related to strategy adoption

David Menendez [*]

Department of Psychology, University of Wisconsin – Madison, Madison, Wisconsin, United States

Sarah A. Brown

Department of Psychological & Brain Sciences, Boston University, Boston, Massachusetts, United States

&

Martha W. Alibali

Department of Psychology, University of Wisconsin – Madison, Madison, Wisconsin, United States

[*] Corresponding author. Email: dmenen@umich.edu. Mailing: 530 Church St, Ann Arbor, MI 48109

Acknowledgements

David Menendez is now at the University of Michigan. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Award #R305B150003 to the University of Wisconsin-Madison. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education. We would like to thank Kylie Robinson, Gill-Helene Schomaker, Nicole Koshevatskiy, and Jaclyn Psenka for their help with coding and data collection.

Some correct strategies are better than others:

Individual differences in strategy evaluations are related to strategy adoption

Abstract

Why do people shift their strategies for solving problems? Past work has focused on the roles of contextual and individual factors in explaining whether people adopt new strategies when they are exposed to them. In this study, we examined a factor not considered in prior work: people's evaluations of the strategies themselves. We presented undergraduate participants from a moderately selective university (*N* = 252; 64.8% women, 65.6% White, 67.6% who had taken calculus) with two strategies for solving algebraic word problems and asked them to rate these strategies and their own strategy on a variety of dimensions. Participants' ratings loaded onto two factors, which we label *quality* and *difficulty*. Participants' initial evaluations of the *quality* of the strategies were associated with whether they used the strategies at posttest, and this effect held even when controlling for individual and contextual factors. However, people's evaluations of the *difficulty* of the strategies were not consistently associated with their later adoption of those strategies. We also examined individual and contextual predictors of strategy ratings and strategy adoption. Participants' need for cognition and their spatial visualization ability were associated with their strategy evaluations, and the framing of the story problems also influenced their strategy adoption. The findings highlight that strategy adoption depends on multiple interacting factors, and that to understand strategy change, it is critical to examine how people evaluate strategies..

Keywords: strategy change; problem solving; individual differences; mathematics learning; strategy ratings

Some correct strategies are better than others: Individual differences in strategy evaluations are related to strategy adoption

1. Introduction

Why do people change their strategies for solving problems? This question is important, as understanding strategy change is critical to understanding cognition, development, and education. As children develop, they often shift from using incorrect or inefficient strategies to using correct or more efficient ones (Siegler, 1996; 2000), and helping students make this shift is a common goal of instruction (Brown & Alibali, 2018a; Fazio et al., 2016; van der Ven et al., 2012). However, people often resist changing their strategies, as extensive research in cognitive psychology has amply demonstrated (Adamson, 1952; Duncker, 1945; Luchins, 1942; McNeil, 2014).

Strategy change is a complex process that involves many individual and contextual factors (Alibali et al., 2019). In this study, we focus on a factor that has been largely neglected in prior research: people's evaluations of the strategies themselves. Specifically, we investigate whether people's evaluations of strategies are associated with their later use of those strategies.

1.1. Strategy adoption

Although people sometimes shift from exclusively using a single strategy to exclusively using a different strategy for a given type of problem (Alibali, 1999), it is more common that people have a repertoire of multiple strategies that they consider and use (Siegler, 1996, 2000). From this perspective, strategy change involves shifts in the set and distribution of strategies that people use from one time point to a later time point. But where do the strategies in that set come from? That is, how do new strategies enter people's strategy repertoires? One possible way is by inventing or discovering new strategies. People can combine elements of strategies they already know in order to create new strategies (Siegler & Jenkins, 1989), or they can notice new features of problems and construct new strategies that rely on those problem features (Alibali et al., 2018). Strategy discovery of this sort sometimes occurs spontaneously, but it is relatively rare (Alibali et al., 2018). However, with explicit prompting to use different strategies, many students can discover new strategies. For example, Star and Rittle-Johnson (2008) prompted middle-school students who had not received instruction on equation solving to generate multiple strategies for solving algebraic equations, and many students were able to do so.

Another way for people to expand their strategy repertoires is by adopting strategies that they encounter in their environment. People may encounter new strategies either via direct instruction or by observing other people's strategies (e.g., in settings that allow for collaborative problem solving; Gutierrez et al., 2018). Learners sometimes encounter *multiple* novel strategies for solving a problem—for example, in a classroom setting in which multiple students are asked to share their approaches to solving a given problem. Past research has shown that learners who are exposed to multiple alternative strategies are more likely to shift their strategy use and to adopt new strategies than learners who are not exposed to alternatives (Brown & Alibali, 2018a; Star & Rittle-Johnson, 2008). Past work further suggests that exposing learners to multiple strategies can lead them to use more efficient strategies, rather than less efficient ones (Star & Rittle-Johnson, 2008). However, people do not always adopt the strategies to which they are exposed (Brown et al., 2019).

If exposure to new strategies does not always lead to strategy adoption, what factors determine whether and when people adopt new strategies? And when faced with multiple novel strategies, how do people choose which one to adopt? Alibali et al. (2019) have argued that strategy adoption depends on multiple factors, including characteristics of the strategy and characteristics of the learner.

1.2. Characteristics of the strategy

1.2.1. Correctness. Past research suggests that learners are sensitive to how often different strategies lead to correct answers. Multiple theories of strategy change, such as strategy selection learning theory (Rieskamp & Otto, 2006) and the RCCL (Represent-Construct-Choose-Learn) model (Lovett & Schunn, 1999), have suggested that people track the success rates of different strategies and shift towards strategies that have higher success rates. These theories can explain why people abandon strategies that yield incorrect solutions and adopt ones that lead to correct solutions. However, this form of associative learning—based on success alone—cannot explain how people choose among *novel* strategies. When people encounter multiple novel strategies, they have no experience with any of them, so all have the same (uninformative) prior success rate (i.e., 0 successes and 0 failures). If learners use only information about success to select strategies, at the first exposure, different novel strategies should be treated equally and adopted at similar rates. However, empirical data show this is not the case. For example, Brown and Alibali (2018a) presented learners with a set of correct and incorrect strategies, but did not tell them which ones were correct. They found that learners were more likely to adopt the correct strategies than the incorrect ones.

When learners have no past experience to draw upon, they may choose a strategy that they believe will get them closer to their goal. Indeed, some strategy-choice models suggest that people evaluate strategies based on their alignment with the goal of the problem at hand. For example, Siegler and colleagues proposed that people apply “goal sketch filters”, which include information about goals and causal relations within the problem domain, when they evaluate potential strategies (e.g., Shrager & Siegler, 1998; Siegler & Araya, 2005; Siegler & Crowley, 1994). Strategies that align with the goal of the problem at hand are allowed through the filter, but strategies that do not align with that goal are filtered out. For instance, the goal sketch filter for a simple addition problem would allow through any novel strategy that uses both addends, but it would filter out one that uses one of the addends twice. In situations in which learners are exposed to both correct and incorrect strategies, but are not told which is which, the learners' goal sketch filters may lead them to adopt a strategy that appears more consistent with the structure of the problem domain. Theories that incorporate such filters can account for why people often avoid adopting strategies that lead to incorrect answers (because they are inconsistent with the goal sketch, so they are not allowed through the filter), and why they prefer strategies that lead to correct answers (because they are consistent with the goal sketch, so they are allowed through the filter).

Such theories, however, cannot explain differences in adoption of different correct strategies, as all such strategies yield correct answers and "pass" the goal sketch filter. Brown et al. (2019) exposed undergraduate students to different sets of strategies for solving word problems. In one condition, students were shown two novel correct strategies, but were not given any other information about the strategies. Given that both strategies were novel, the prior experienced "success rate" for each strategy was the same. Further, because both strategies were correct, both were aligned with the problems' goals. Thus, the aforementioned models of strategy change would predict that participants should have adopted the two strategies at similar rates (Lovett & Schunn, 1999; Shrager & Siegler, 1998; Siegler & Araya, 2005; Siegler & Shipley, 1995; Rieskamp & Otto, 2006). However, Brown et al. (2019) found that this was not the case. In their study, the two strategies to which participants were exposed were an arithmetic strategy and a geometric strategy for solving algebraic word problems. Surprisingly, participants adopted one strategy twice as often as the other! Given that the participants had no information about the prior success of the strategies or about their correctness, they must have relied on strategy preferences that were based on something beyond simple correctness.

1.2.2. *Strategy evaluations.* A few studies have investigated how people evaluate problem-solving strategies. For example, Siegler and Crowley (1994) asked children to judge strategies (both strategies for solving arithmetic problems and strategies for playing tic-tac-toe) as *smart*, *kind of smart*, or *not so smart*. They found that children judged correct strategies as smarter than incorrect strategies. However, strategies differ in many dimensions beyond whether they are correct or not. For example, some strategies have fewer steps than others, and some strategies might be easier to understand. Brown et al. (2018) considered multiple dimensions in undergraduates' evaluations of three correct strategies for solving algebraic word problems. They found that undergraduates' ratings could be explained by two factors, which they termed *intuitiveness* and *efficiency*. The *intuitiveness* factor included ratings on items such as: "how common is this strategy?", "how good is this strategy?", and "how much sense does this strategy make?". The *efficiency* factor included ratings on items such as "how complicated is this strategy?", "how easy is this strategy to remember?", and "how long would this strategy take?" Brown et al. (2018) found that participants' ratings of the three strategies varied, but they did not provide any evidence that these ratings played a role in whether participants adopted the strategies. In this study, we examine whether people's ratings about strategies predict which strategies they adopt.

1.3. Learner characteristics

Past research has shown that some characteristics of learners also predict their likelihood of adopting strategies. Here we focus on three characteristics that have been considered in prior work: confidence in their prior strategy, need for cognition, and spatial visualization ability. We also explore individuals' strategy preferences, a factor that has received little attention in prior work.

1.3.1. Confidence. Prior work has shown that people's confidence—defined as "feeling of success (predicted or achieved) in a task"— influences their decision making (Aguilar-Lleyda et al., 2020, p. 1084). Confidence might be used as signal of correctness when there is no feedback (Guggenmos et al., 2016; Hainguerlot et al., 2018). In the context of adopting new strategies, learners who are very confident that their current strategy is correct are less likely to adopt a new strategy than those who lack confidence in their current strategy (Brown et al., 2019).

1.3.2. Need for cognition. Need for cognition is the tendency to engage in and enjoy complex, effortful cognitive activity (Cacioppo & Petty, 1982; Sadowski & Gülgöz, 1992). Prior work has found that people high in need for cognition are more likely to adopt a novel strategy after being exposed to it than people low in need for cognition (Brown et al., 2019). Some work has further suggested that need for cognition interacts with confidence to predict strategy change, in that need for cognition matters less when participants are very confident that their original strategy is correct (Brown et al., 2019). However, some other work has failed to replicate this interaction (Brown & Alibali, 2018a). Additionally, need for cognition might be related to how participants evaluate strategies, as people high in need for cognition might think more deeply about possible strategies and why they work (Brown et al., 2018). Finally, need for cognition may be particularly important for adopting certain strategies. Prior work has found that the effect of need for cognition on strategy adoption is stronger for difficult or unintuitive strategies than for simpler, more common strategies (Brown et al., 2019).

1.3.3. Spatial visualization ability. Several studies have shown a link between spatial ability and achievement in mathematics (Hegarty & Kozhevnikov, 1999; Uttal et al., 2013; Wai et al., 2009). In this work, we focus on one specific aspect of spatial ability, the ability to mentally visualize and transform objects, which we term *spatial visualization ability*. In the taxonomy of spatial ability offered by Newcombe and Shipley (2015; see also Newcombe, 2018), this ability is considered a form of intrinsic-dynamic spatial ability. We focus on this aspect of spatial ability for two reasons. First, intrinsic-dynamic spatial ability is associated with successful mathematical problem solving (Lubienski et al., 2021). Second, spatial visualization ability may relate to people's evaluations of strategies for solving mathematical problems that involve visual representations, given past studies showing spatial visualization ability is related to how people engage with and learn from visual representations (Bartel & Alibali, 2021; Bartholomé & Bromme, 2009; Hegarty, 2011; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997).

1.3.4. Strategy preferences. Prior work on strategy use in chemistry education has shown that people sometimes have preferences for certain types of strategies (e.g., diagrammatic strategies or algorithmic strategies; Stieff et al., 2012), but there is little work on people's preferences for mathematical strategies. People may value different characteristics of strategies. For example, some people might prefer strategies that have as few steps as possible, while others might value strategies that are intuitive and easy to understand. In this study, we take a first step towards examining whether such general strategy preferences influence strategy adoption.

1.4. Current study

The main goal of this study was to examine whether participants' evaluations of strategies predicted their subsequent use of those strategies to solve problems. To address this goal, we exposed undergraduate students to two correct strategies for solving an algebraic word problem. Participants rated each strategy on a variety of dimensions, and then solved similar problems.

We also considered whether the two target strategies were adopted at differential rates. Given that the two strategies were both novel and correct, existing models of strategy change would suggest that participants should be similarly likely to adopt the two strategies. However, past research has shown that participants are more likely to adopt some strategies than others (e.g., Brown & Alibali, 2018). Therefore, we examined whether rates of strategy adoption varied for the two target strategies, and whether adoption depended on problem features.

We also sought to replicate past findings on individual characteristics as predictors of strategy adoption. As reviewed above, past work has identified several characteristics of learners that predict adoption of novel strategies, including high need for cognition and low confidence in existing strategies. We considered associations of these individual difference factors with strategy adoption, as well.

1.4.1. Task domain: Constant change problems

We examined strategy change in undergraduates solving constant change problems. Constant change problems are algebraic word problems that describe a rate that changes over a given interval of time or space. For example, one of the problems used in the study was: "Milk is pumped into a vat for a period of 12 minutes. The rate at which it is pumped increases steadily over the interval from 7 gallons per minute to 139 gallons per minute. How many gallons are pumped into the vat over the 12-minute interval?"

Constant change problems can focus either on quantities that change continuously (e.g., milk being pumped continuously into the vat) or quantities that change discretely (e.g., books on a bookshelf, with the number of books on each subsequent shelf increasing by a constant number). We refer to problems about quantities that change continuously as *continuously-framed* and problems about quantities that change discretely as *discretely-framed*. Although the underlying mathematics of the problems is the same, and all correct strategies work for both continuous and discrete problems, past work has shown that solvers often conceptualize continuous and discrete problems differently (Brown & Alibali, 2018b). Problem wording can also be used to cue continuous and discrete representations of the problems (e.g., Alibali et al., 1999; Brown & Alibali, 2018b).

This study builds on prior research that identified and examined multiple strategies for solving constant change problems (Alibali et al., 1999; Brown et al., 2018; 2019; Riggs et al., 2015; 2017). In these previous studies, the most common strategy used by undergraduates was the *summation strategy*. For the milk pumping problem, people who use the summation strategy calculate how many gallons are pumped into the vat in each minute and then sum these values to find a total. In the less common *Gauss strategy* (named for the mathematician Carl Friedrich Gauss, who purportedly invented the strategy), people add the number of gallons pumped in the first minute and the last minute and multiply that sum by the number of minutes divided by two. Another less common approach was the *area strategy*, in which people draw a visual representation of the problem and calculate the area, which corresponds to the total number of gallons, using formulas for areas of shapes. In rare cases, participants calculated the area using integration. Prior research has shown that when exposed to both the area strategy and the Gauss strategy, undergraduates adopt the Gauss strategy more frequently (Brown, et al., 2019).

Undergraduates also provide more positive evaluations of the Gauss strategy than the area strategy (Brown et al., 2018). It is worth noting that the Gauss strategy has only two steps, whereas the area strategy has three steps. Additionally, undergraduates are more likely to use the summation strategy on discretely-framed problems than on continuously-framed problems (Brown & Alibali, 2018b).

We presented the strategies using worked examples that showed a series of steps that could be used to solve a given problem. Prior work has shown that worked examples can enhance learning (Atkinson et al., 2000; Booth et al, 2015; Durkin et al., 2021) and increase adoption of problem-solving strategies (Star & Rittle-Johnson, 2008; Barbieri & Booth, 2016). Worked examples reduce learners' cognitive load by allowing them to focus on learning how to solve the problem, rather than on actually solving the problem (Paas & van Merriënboer, 1994).

1.4.2 Research questions and hypotheses

In this study, our primary research question concerned whether participants' evaluations of the strategies predicted their subsequent use of those strategies. In this regard, we were especially interested in whether participants' strategy evaluations would predict strategy adoption, over and above individual characteristics and baseline rates of adoption for each strategy. It is sensible to expect that participants' evaluations of the strategies would predict strategy adoption; however, given that there is no prior work examining this relation, we did not pre-register specific hypotheses regarding the relation between strategy evaluations and strategy adoption.

However, given past work, we did make specific predictions regarding individual characteristics and strategy effects. Our pre-registered hypotheses were: (1) participants would be more likely to adopt the Gauss strategy than the area strategy; (2) participants who scored higher on a Need for Cognition scale would be more likely to adopt a new strategy (at least for the less intuitive area strategy), and (3) participants who were more confident about their pretest strategy would be less likely to adopt a new strategy. In addition, based on prior research (Brown et al., 2019), we expected that, for participants who had low confidence in their pretest strategy, their level of need for cognition would not be associated with adoption of the Gauss strategy, but it would be associated with adoption of the area strategy. Therefore, we included the interaction of pretest confidence and need for cognition in each of the models that explored the role of individual differences in adoption of each strategy. The pre-registration for this study can be found at this link: https://osf.io/mj68b/?view_only=2a842125d3dc4ea79f91bdde9bace6bf

In addition to these pre-registered hypotheses, we also explored some other individual difference factors for which we did not advance specific hypotheses. Because one of the target strategies included a visual representation, we considered whether individual differences in spatial visualization ability would predict adoption of that strategy. In addition, given our interests in strategy evaluations, we considered whether general preferences for certain types of strategies (e.g., preferences for short-cuts) would predict strategy adoption. Finally, we also explored whether individual characteristics were associated with patterns of strategy evaluations.

2. Method

2.1. Participants

Participants were 252 undergraduate students who were enrolled in an Introduction to Psychology course at a moderately selective (58% acceptance rate), large Midwestern University. They received extra credit in the course for participating in the study. Two participants were excluded because the experimenter accidentally gave them additional problems to solve. Due to this experimenter error, we excluded these two participants from our analyses, even though we did not pre-register this exclusion criterion. Demographic information is presented in Table 1.

Table 1. Participant demographic information.

		Frequency (%)
Gender		
	Women	162 (64.8%)
	Men	88 (35.2%)
Race/ethnicity		
	White	164 (65.6%)
	Asian or Asian American	50 (20.0%)
	Hispanic or Latinx	7 (2.8%)
	Black or African American	6 (2.4%)
	Native American	2 (0.8%)
	Middle Eastern	2 (0.8%)
	Bi- or multi-racial	18 (7.2%)
	Did not disclose	1 (0.4%)
Year in school		
	First	211 (84.4%)
	Second	28 (11.2%)
	Third	8 (3.2%)
	Fourth	3 (1.2%)
Highest level of prior or concurrent mathematics coursework		
	Geometry, Algebra, or pre-calculus	78 (31.2%)
	One semester of calculus	63 (25.2%)

Two semesters of calculus	78 (31.2%)
More than two semesters of calculus	28 (11.2%)
Statistics	3 (1.2%)
Mean standardized math score (SD)	88.40 (11.61)
Average Need for Cognition score across the 18 items (out of 5) (SD)	3.10 (0.56)
Mean score (number correct) on the Paper Folding task (out of 20) (SD)	11.75 (3.46)

2.2. Materials

A list of the problems used in the study can be found at: https://osf.io/m6pyv/?view_only=667f55ef071c47de861851b53723bcc0.

2.2.1. *Pretest.* The pretest consisted of one continuously-framed constant change problem. Participants were given 5 minutes to solve the problem, and after solving it, they rated how confident they were that they solved it correctly on a 1 (I am sure I did it wrong) to 5 (I am sure I did it right) scale. Then, participants rated the strategy they used to solve the problem using a 1 (not at all X) to 5 (very X) scale on the following dimensions:

1. How good is your strategy?
2. How common is it for people to use your strategy to solve this kind of problem?
3. How complicated is your strategy? (reverse coded for analysis)
4. How easy would it be to remember your strategy?
5. How long did it take to use your strategy? (reverse coded for analysis)
6. Does your strategy make sense?

7. How efficient is your strategy?
8. How intuitive is your strategy?

2.2.2. *Exposure*. In the exposure phase, participants saw two continuously-framed constant change problems, each accompanied by an explanation of how a student solved the problem. Participants received no information about the students whose strategies they saw. For one of the problems, participants saw a worked example of the area strategy, and for the other they saw a worked example of the Gauss strategy. The order in which the strategies were presented and the problem that accompanied the strategy were counterbalanced. After reading the worked example of each strategy, participants rated the strategy using the same 8-item scale that they had used to rate their own strategy, with the questions modified to refer to "this strategy" rather than "your strategy." This scale had good internal consistency (Cronbach's $alpha_{Gauss} = 0.83$; $alpha_{area} = 0.87$; $alpha_{pretest} = 0.83$). Participants were also asked, "How likely do you think it is for you to get the correct answer when using this strategy?", which was analogous to the confidence question that was asked at pretest.

2.2.3. *Posttest*. The posttest consisted of two constant change problems. The first was continuously framed, and the second was discretely framed. Given that the novel strategies had been presented with a continuously-framed problem, the use of the novel strategy on the discretely-framed posttest problem served as an indicator of generalization of the novel strategy. This is a stringent test of generalization, as prior work suggests that people are less likely to use the Gauss and area strategies for discretely-framed problems than for continuously-framed problems (Alibali & Booth, 2002; Brown & Alibali, 2018b).

2.2.4. *Individual difference survey.* Participants completed the individual difference measures on a computer. These measures were: (1) the Paper Folding Test, (2) the Need for

Cognition scale, (3) a set of general strategy preference questions, and (4) a demographic questionnaire, which included questions about mathematics ability and experience.

2.2.4.1. Paper Folding Test. Participants completed a computerized version of the Paper Folding Test (Ekstrom et al., 1976) to measure their spatial visualization ability, which is a form of intrinsic-dynamic spatial ability (Newcombe & Shipley, 2015). On each of the 20 trials of this test, participants are shown a drawing of piece of paper that has been folded multiple times. After the folds, the paper is punctured, creating a set of holes. On each trial, participants view five options and must select the one that shows how the paper would look if unfolded. For our sample, Cronbach's alpha for this test was 0.73.

2.2.4.2. Need for Cognition scale. We used the short form of the Need for Cognition scale (Cacioppo et al., 1984). This measure consists of 18 statements which are rated on scale from 1 (extremely uncharacteristic of me) to 5 (extremely characteristic of me). Some example statements include "I would prefer complex to simple problems," "Thinking is not my idea of fun" (reverse coded), and "I really enjoy a task that involves coming up with new solutions to problems." For our sample, Cronbach's alpha for this test was 0.86.

2.2.4.3. Strategy preference questions. This measure was not described in the pre-registration, and thus, its inclusion is a deviation from our pre-registered protocol. We asked participants three questions about their general preferences for problem-solving strategies. The three questions were: "In general, when solving problems, I like to use shortcuts even when I don't know how they work," "In general, when solving problems, I prefer to use strategies that I understand well," and, "In general, when solving problems, I prefer to use strategies that have fewer steps." Participants answered these questions on a scale from 1 (extremely uncharacteristic of me) to 5 (extremely characteristic of me). We analyzed responses for each question separately, so we did not calculate Cronbach's alpha for these items.

2.2.4.4. Demographics. Participants completed a demographic questionnaire that requested information about their previous math coursework, SAT and/or ACT math scores, year in college, gender, age, and race/ethnicity.

2.3. Procedure

Participants completed the study in a computer lab. Participants first received the pretest problem and were given 5 minutes to complete it using pen and paper. If they finished the problem early, they were asked to wait in their seat for the 5 minutes to pass. Participants were then given the exposure packet. They were asked to read the strategies and complete the rating scales. When they were done, they received the posttest problems and were given up to 30 minutes to complete them, again with pen and paper. All participants finished within the allotted time. When each participant was finished with the posttest, they completed the individual difference survey at a computer.

2.4. Strategy coding

For each problem, we coded whether participants used summation, area, Gauss, or some other strategy. Strategies categorized as "other" were primarily incorrect (e.g., subtracting the initial from the final rate) and idiosyncratic. A small subset of the strategies categorized as "other" involved attempts to use integration, and most of these attempts were incorrect. Participants could receive credit for using multiple strategies on one problem. One trained coder coded the pretest and posttest strategies for all participants. A second coder independently coded the pretest and posttest strategies for 69 participants (27.6% of the sample). We calculated Cohen's kappa for each category, and reliability was acceptable for all categories: summation (κ

= 0.85), area (κ = 0.84), Gauss (κ = 0.71), and other (κ = 0.72). All disagreements on the reliability sample were resolved through discussion, and the agreed-upon codes were used in the final analyses.

2.5. Transforming standardized mathematics scores

We transformed participants' self-reported ACT or SAT math scores into standardized math scores using percentile conversion tables from each participant's high school senior year, as inferred from their reported year in college. If the participant reported both ACT and SAT math scores, we used the higher percentile score. Some participants reported their SAT combined score; for these participants, we used the percentile of their combined score, if their ACT math score was not available. Seven participants had missing data for standardized mathematics scores, and their data was excluded from the analyses.

3. Results

3.1. Analysis plan

We first present analyses of the strategies that participants used at pretest and posttest. These analyses show that exposure to new strategies can lead to strategy change and that strategy adoption depends on problem features. We used chi-square tests to examine whether the distribution of strategies differed for discretely-framed and continuously-framed problems. We then present analyses of how participants rated the different strategies. To control for type I error rate, we performed an omnibus test to examine whether there were differences in ratings by strategy, and we performed pairwise comparisons only if this test was significant.

Next, we present our pre-registered analyses examining the factor structure of the strategy ratings. We first used confirmatory factor analysis (CFA) in an attempt to replicate the findings from Brown et al. (2018b) using only the ratings included in their study. Then we performed the pre-registered CFA using all the ratings. These CFA did not fit the data well, sowe deviated from out pre-registered plan and conducted an exploratory factor analysis. Using the results of this EFA, we computed factor scores for each participant using the Thurstone method as specified by Grice (2001).

We then used these factor scores in our analyses of whether strategy ratings and individual differences were associated with strategy adoption. We report the likelihood that participants adopted each of the two target strategies (Gauss and area) on the first, continuously-framed posttest problem and the likelihood that participants generalized each strategy to the second, discretely-framed posttest problem. Therefore, we fit four logistic models one predicting adoption of area on the first problem, one predicting adoption of Gauss on the first problem, one predicting adoption of area on the second problem, and one predicting adoption of Gauss on the second problem. As predictors, we included participants' ratings of quality and difficulty for that strategy, scores on the Paper Folding Test, standardized mathematics scores, Need for Cognition scores (mean-centered), confidence in their pretest strategy (mean-centered), and the interaction between confidence in their pretest strategy and Need for Cognition scores. Finally, we present exploratory analyses of how individual differences are associated with strategy ratings.

3.2. Strategy Use

As in prior research (Brown et al., 2019; Riggs et al., 2015, 2017), the majority of participants used the summation strategy (56.8%) or a strategy classified in the "other" category (46.0%) at pretest. Use of the area strategy (0.4%) and the Gauss strategy (2.8%) were extremely rare at pretest. As shown in Figure 1, the distribution of strategies for the first posttest problem, which was continuously framed, differed from the distribution of strategies for the pretest problem, which was also continuously framed, $\chi^2(3, N = 250) = 175.80$, $p < .001$. Many participants used summation on the continuous posttest problem (29.2%), but many more used the area (22.8%) and Gauss (38.8%) strategies on the continuously-framed posttest problem than had used it on the pretest problem. Thus, many participants adopted the strategies to which they were exposed, and they used those strategies on the posttest problem that was similar to the exposure problem.

We also examined how the distribution of strategies varied between the two posttest problems. The distribution of strategies differed for the continuously-framed posttest problem and the discretely-framed posttest problem, $\chi^2(3, N = 250) = 48.86$, $p < .001$. As can be seen in Figure 1, many participants (61.6%) used the summation strategy and fewer used the Gauss (26.0%) and area (9.6%) strategies on the discretely-framed posttest problem than on the continuously-framed posttest problem. Thus, not all participants generalized the Gauss strategy to the discretely-framed problem, and even fewer generalized the area strategy.

To test whether participants were more likely to adopt the Gauss strategy than the area strategy, we fit a mixed-effects logistic regression predicting whether participants ever used the strategies on either of the two posttest problems. We included strategy (Gauss or area) as a predictor, as well as by-participant random intercepts and by-participant random slopes for the effect of strategy. As hypothesized, and replicating Brown et al. (2019), participants were more likely to adopt the Gauss strategy than the area strategy, $OR = 2.90$, $\chi^2(1, N = 250) = 8.35$, $p = .004$.

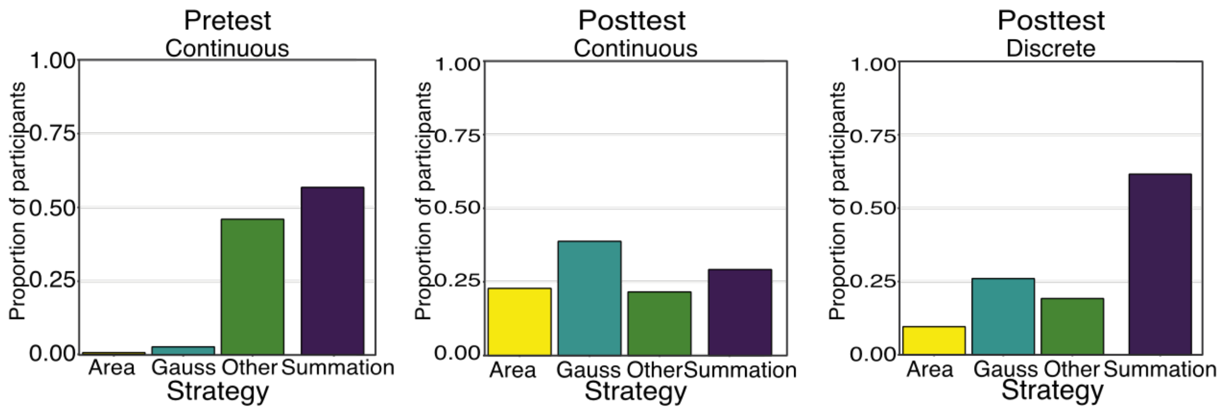

Figure 1. Distribution of strategies at pretest and posttest. The y-axis shows the proportion of participants who used each strategy. The first panel shows the distribution for the pretest problem, which was continuously framed. The middle panel shows the distribution for the first posttest problem, which was also continuously framed. The last panel shows the distribution for the second posttest problem, which was discretely framed.

3.3. Strategy ratings

We next examined participants' ratings of the strategies. Given the high Cronbach's alpha for the scale as a whole, we first averaged the ratings of the eight dimensions (after reverse coding ratings for how long and how complicated the strategy was) for each participant and each strategy. We fit a linear mixed-effects model predicting participants' ratings with strategy as a predictor (area, Gauss, and pretest, with the pretest strategy as the reference group). We included by-participant random intercepts and by-participant random slopes for the strategy contrasts, and we allowed the random effects to correlate, but this model did not converge. We followed recommendations by Brauer and Curtin (2018) to achieve convergence. The first model to converge did not allow the random effects to correlate. Overall, participants' ratings differed across the strategies, $F(2, 358.19) = 15.16$, $p < .001$. On average, participants rated the Gauss strategy more favorably ($M = 3.44$, $SD = 0.68$) than their pretest strategy ($M = 3.26$, $SD = 0.77$),

$F(1, 352.13) = 7.70$, $p = .006$, and they rated their pretest strategy more favorably than the area strategy ($M = 3.08$, $SD = 0.77$), $F(1, 361.83) = 7.91$, $p = .005$.

We fit the same model for ratings of each of the dimensions. There were differences among the strategies in ratings of goodness ($F(2, 351.55) = 19.21$, $p < .001$), commonness ($F(2, 371.64) = 27.90$, $p < .001$), complexity ($F(2, 350.09) = 83.41$, $p < .001$), easiness to remember ($F(2, 358.36) = 25.86$, $p < .001$), length ($F(2, 371.56) = 62.10$, $p < .001$), intuitiveness ($F(2, 365.18) = 13.01$, $p < .001$), and efficiency ($F(2, 358.71) = 28.55$, $p < .001$). Differences in whether the strategies made sense were not significant, $F(2, 349.86) = 2.43$, $p = .089$. Figure 2 presents the ratings for each strategy for each dimension and indicates the results of pairwise comparisons between the strategies. On the whole, the pairwise comparisons suggest that participants did not “default” to rating their pretest strategy as better than the alternatives. Although, on average, participants rated their pretest strategy as the most common, easiest to remember, and least complex, they also rated it as the least good, least efficient, and least intuitive. Participants also differed in their evaluations of the two novel strategies. Specifically, participants rated the Gauss strategy as more common, easier to remember, shorter, and more efficient than the area strategy.

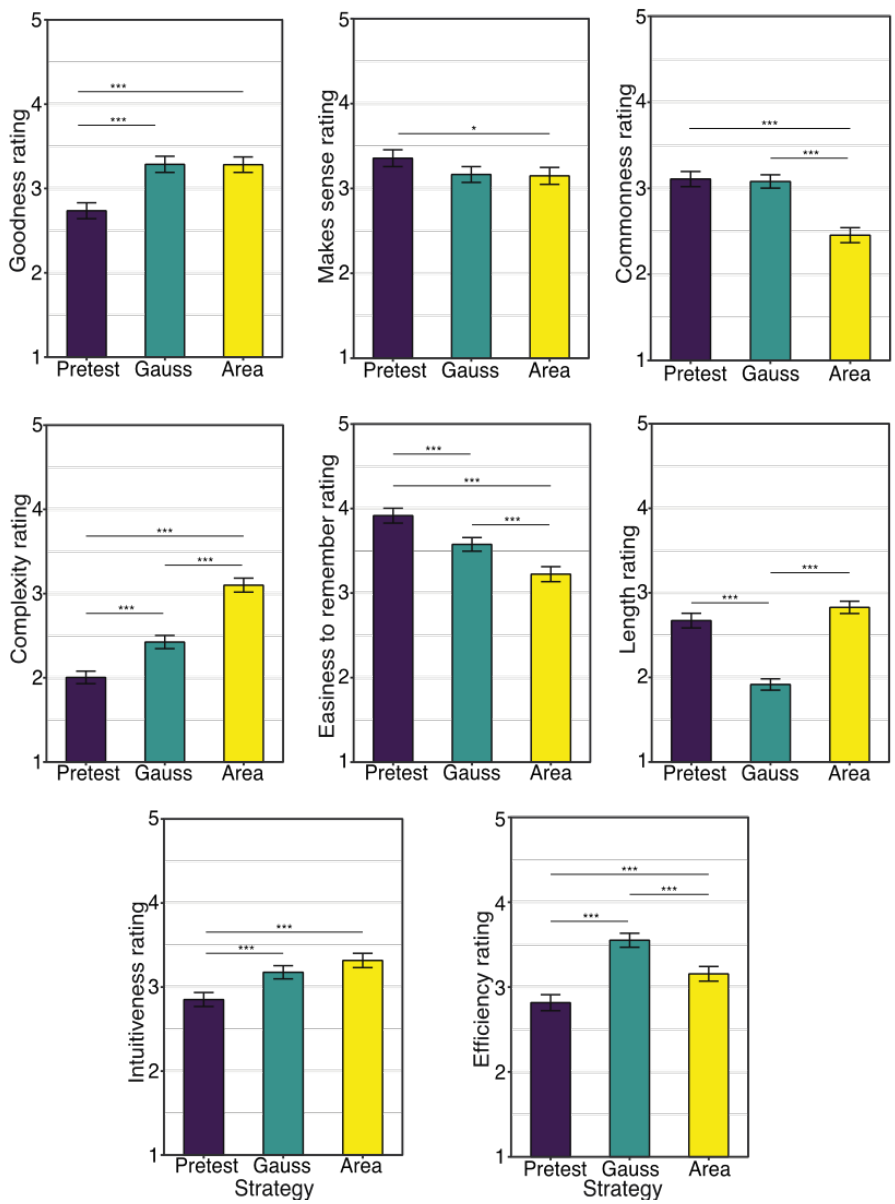

Figure 2. Mean ratings of goodness, making sense, commonness, complexity, easiness to remember, length, intuitiveness, and efficiency for each strategy. Error bars show the within-participant standard errors. * $p < .05$ *** $p < .001$

3.4. Pre-registered analyses

In our pre-registration, we specified that we would exclude participants who did not use summation at pretest, so that our analytic sample would be comparable to that used in prior research (Brown et al., 2019; Riggs et al., 2015, 2017). Of the 250 participants, 142 used summation at pretest. We present the analyses for this subsample; except where noted, the results are unchanged if we include the full sample (see Supplemental materials). First, we analyzed the data using the factor structure presented by Brown et al. (2018), which used items that tapped six dimensions: commonness, goodness, making sense, complexity, easiness to remember, and length. Then, we present our pre-registered factor analysis, which also includes items in which participants rated intuitiveness and efficiency.

3.4.1. Analysis attempting to replicate Brown et al.'s (2018b) factor structure. For each strategy, we averaged participants' ratings for items hypothesized to load on each dimension, and used these average ratings in our factor analysis. In Brown et al.'s (2018) factor analysis, commonness, goodness, and making sense loaded onto an "intuitiveness" factor, and complexity, easiness to remember, and length loaded onto an "efficiency" factor. This two-factor solution did not fit the current data well, $\chi^2(8, N = 142) = 98.85$, $p < .001$, BIC = 3568.48, CFI = .811, RMSEA = .204, 90% CI [.169, .241]. However, a single factor model also did not fit the data well, $\chi^2(9, N = 142) = 103.45$, $p < .001$, BIC = 3568.48, CFI = .671, RMSEA = .254, 90% CI [.211, .299]. The two-factor model had a lower BIC, a lower RMSEA, and a higher CFI than the single-factor model.

3.4.2. Pre-registered factor analysis. We then tested the pre-registered factor analysis (which included the ratings of intuitiveness and efficiency, which had not been included in the Brown et al. [2018b] study). In this hypothesized model, there are two factors: intuitiveness (made up of intuitiveness, commonness, goodness, and making sense) and efficiency (made up of efficiency, complexity, easiness to remember, and length). This hypothesized model did not fit the data well $\chi^2(19, N = 142) = 123.27$, $p < .001$, BIC = 4565.70, CFI = .790, RMSEA = .197, 90% CI [.164, .230]. A single-factor model also did not fit the data well, $\chi^2(20, N = 142) =$

126.66, $p < .001$, BIC = 4563.01, CFI = .788, RMSEA = .193, 90% CI [.162, .226]. The fit indices suggested that the single-factor model was preferred.

Given that the hypothesized model was not supported, we deviated from our pre-registered analysis plan and conducted an exploratory factor analysis. We determined the number of factors to extract in two ways. First, we examined the scree plot of successive eigenvalues and looked at the number of items before the elbow. The scree plot (Figure 3) suggested that we should extract two factors. Second, we fit models extracting between one and four factors, and we selected the model with the lowest BIC as the best model. We conducted this exploratory factor analyses using a varimax rotation and maximum likelihood extraction. The BICs also indicated that the model with two factors was the best model (1 factor: BIC = 23.2, 2 factors: -28.8, 3 factors = -19.5, 4 factors = -5.9), and it was an acceptable fit for the data, TLI = .90, RMSEA = 0.11, 90% CI [0.068, 0.155]. We used factor loadings greater than 0.40 as the cutoff for whether an item was included in a factor. See Table 2 for factor loadings. Intuitiveness, commonness, goodness, making sense, and efficiency loaded onto one factor, and complexity, easiness to remember, and length loaded onto the other factor. Note that the only difference between this model and the pre-registered model was that efficiency did not load on the "efficiency" factor, but rather loaded on the "intuitiveness" factor. On this basis, we concluded that the initial names we had given to the factors were not accurate. We suggest that what Brown et al. (2018) termed "intuitiveness" might be better characterized as the perceived *quality* of the strategies, with strategies that are more common, make more sense, are more efficient, are more intuitive, and are perceived as "better" being higher in quality. Further, we suggest that what Brown et al. (2018) termed "efficiency" might be better characterized as the perceived *difficulty* of the strategy, with strategies that are more complex, longer, and less easy to remember being more difficult.

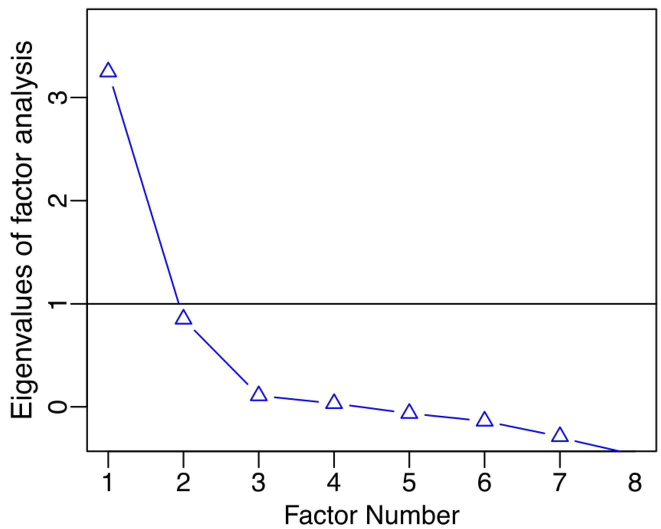

Figure 3. Scree plot showing eigenvalues.

Table 2. Factor loadings for each item for the 2-factor exploratory factor analysis model. Note: all of the factor loadings for Difficulty were reverse scored for ease of interpretation.

	Quality	Difficulty
Goodness	**0.89**	-0.03
Efficiency	**0.80**	-0.15
Makes sense	**0.77**	-0.04
Intuitiveness	**0.80**	-0.24
Commonness	**0.47**	-0.33
Complexity	-0.15	**0.77**
Easiness to remember	-0.33	**-0.62**
Length	0.18	**0.56**

3.4.3. Predicting adoption. As described in our pre-registered analysis plan, we planned to examine whether participants' ratings of strategies predicted their strategy adoption. First, we examined adoption of each of the target strategies on the first posttest problem. This problem was continuously framed, which matched the framing participants saw on the pretest problem and on the problems they saw during the exposure phase. Results are presented in Table 3. For both the Gauss and area strategies, participants who rated the strategy as higher quality were more likely to adopt that strategy. Additionally, participants who rated the area strategy as higher in difficulty were less likely to adopt that strategy. This was also the case for the Gauss strategy, but the relation was not significant (though it was significant in the analysis of the full sample; see the supplemental materials). See Figure 4. As predicted (and replicating Brown et al., 2019), participants were also less likely to adopt the Gauss strategy if they were more confident in their pretest strategy. Participants with higher spatial visualization abilities were also less likely to adopt the Gauss strategy. No other effects were significant. Of note, we did not replicate the interaction of confidence and need for cognition on adoption of the area strategy that was reported by Brown et al. (2019). In the full sample, this interaction was significant, but the pattern differed from that observed in Brown et al. (2019); see the supplemental materials.

Table 3. Results of logistic regressions examining strategy adoption for the first, continuously-framed posttest problem. Values in in **bold** indicate statistically significant results.

	Outcome: Adopting Gauss strategy			Outcome: Adopting area strategy		
	OR	χ^2	*p*	*OR*	χ^2	*p*
Quality	**3.20**	**25.48**	**< .001**	**2.52**	**14.62**	**< .001**
Difficulty	0.62	3.68	.055	**0.46**	**6.83**	**.009**

Need for Cognition	2.07	2.55	.110	1.36	0.44	.509
Confidence	**0.55**	**11.15**	**< .001**	1.13	0.39	.532
Need for Cognition x Confidence	1.29	0.76	.382	1.19	0.35	.553
Spatial visualization ability	**1.21**	**6.99**	**.008**	1.10	1.72	.190
Standardized math score	1.01	0.29	.589	0.98	1.18	.276

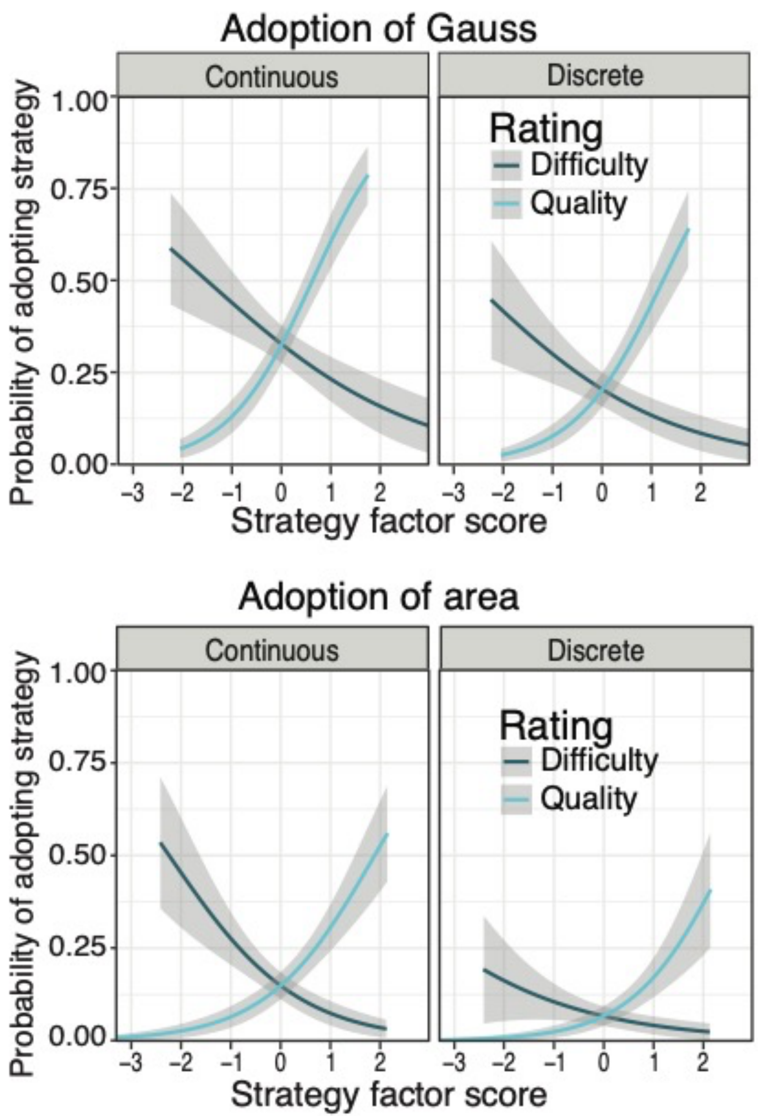

Figure 4. Model predictions showing the relations between ratings of quality and difficulty and strategy adoption for the Gauss and area strategies. The *y*-axis shows the probability of adopting

the given strategy, and the *x*-axis shows the ratings for that strategy. The top two panels show the results for the Gauss strategy, and the bottom two panels show the results for the area strategy. The left panels show the results for the continuously-framed problem (the first posttest problem), and the right panels show the results for the discretely-framed problem (the second posttest problem). Error bands show the within-subject standard errors of the point estimates. Note that the models included ratings of quality and difficulty as two separate predictors, and they do not test for the interaction between these two factors. When the lines start after the value of -3 on the *x*-axis, it is because no participant gave a lower rating.

Next, we examined strategy adoption for the second posttest problem. This problem was discretely framed, which did not match the framing of the pretest or the exposure problems. Additionally, prior research suggests that people frequently use summation for discretely-framed problems (Brown & Alibali, 2018b). Therefore, use of one of the novel strategies on this problem serves as a measure of generalization of the area or Gauss strategy. Once again, for both the Gauss and the area strategies, participants who provided higher quality ratings were more likely to adopt the strategy. Additionally, participants who rated the Gauss strategy as higher in difficulty were less likely to adopt the strategy. This was also the case for the area strategy, but the relation was not significant. In analyses of the full sample, ratings of difficulty were not significantly related to adoption for either strategy.

For this problem, there was also a significant interaction of confidence and need for cognition for adoption of the Gauss strategy. To explore this interaction, we recentered confidence to one standard deviation above and below the mean. As can be seen in Figure 5, participants with low confidence in their pretest strategy were similarly likely to adopt Gauss regardless of their level of need for cognition, *OR* = 1.05, $\chi^2(1, N = 136) = 0.01$, $p = .916$. These participants had little confidence that their prior strategy was correct, so they were moderately likely to try something new, and this tendency did not depend on their level of need for cognition. In contrast, for participants with high confidence in their pretest strategy, higher need for cognition was associated with higher likelihood of adopting the Gauss strategy, *OR* = 5.02, $\chi^2(1, N = 136) = 5.88$, $p = .015$. Put another way, those who were highly confident that their prior strategy was correct were unlikely to try something new, unless they also had high need for cognition. The pattern was similar in the full sample, but the interaction was not significant (see the supplemental materials). It is worth noting that Brown et al. (2019) also found a significant interaction of participants' confidence in their prior strategy and their need for cognition on strategy adoption, but the data pattern differed from that reported here. We consider the differences in our findings and those of Brown et al. (2019) in the discussion.

Additionally, participants with higher standardized math scores were more likely to adopt the Gauss strategy, but standardized math scores were not associated with adoption of the area strategy. No other effects were significant. See Table 4.

Table 4. Results of logistic regressions examining strategy adoption on the second posttest problem (which was discretely framed). Values in in **bold** indicate statistically significant results.

	Outcome: Adopting Gauss strategy			Outcome: Adopting area strategy		
	OR	χ^2	*p*	*OR*	χ^2	*p*
Quality	**3.04**	**22.83**	**< .001**	**2.86**	**11.42**	**< .001**
Difficulty	**0.60**	**3.89**	**.048**	0.61	1.75	.185
Need for Cognition	2.30	3.26	.071	2.10	1.59	.207

Confidence	0.76	1.99	.157	0.72	1.84	.175
Need for Cognition x Confidence	**1.79**	**4.13**	**.042**	1.22	0.30	.581
Spatial visualization ability	1.02	0.05	.825	1.13	1.68	.194
Standardized math score	**1.11**	**8.71**	**.003**	0.97	1.34	.247

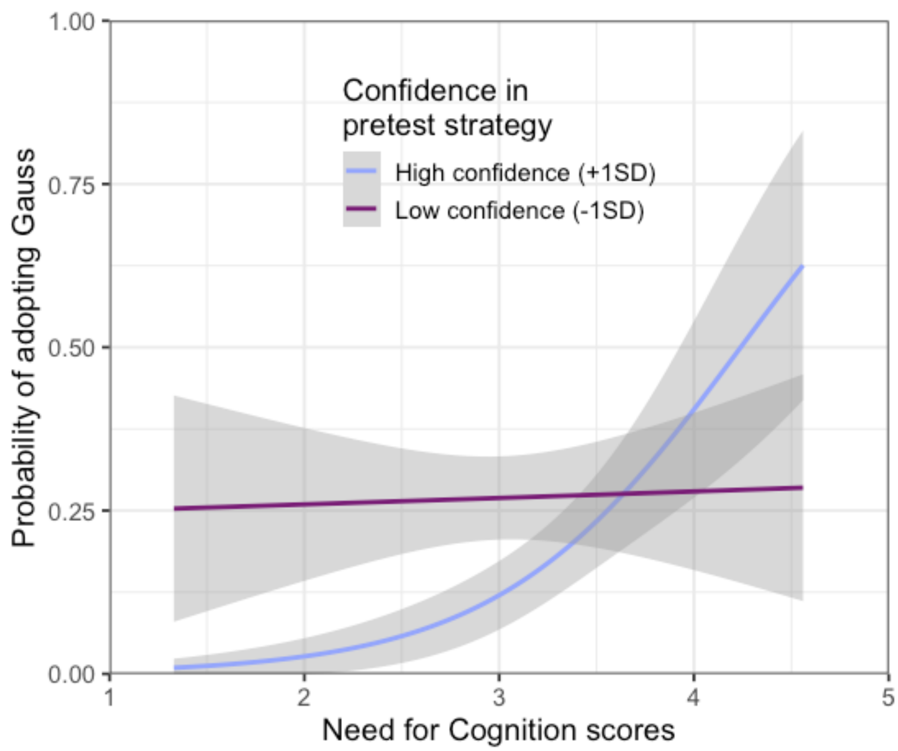

Figure 5. Probability of adopting the Gauss strategy on the discretely-framed posttest problem, as function of Need for Cognition scores (x-axis), and broken down by whether participants had high or low (+/- 1 SD) confidence in their pretest strategy (different lines). The error bands show the within-subject standard errors of the point estimates.

3.5. Exploratory analyses

3.5.1. Individual differences predict ratings. We also explored whether individual difference characteristics predicted participants' ratings of the strategies. For this analysis, we used the full sample (i.e., including those who did not use summation at pretest) in order to increase power, and because we did not know whether strategy use at pretest would be related to participants' ratings. We fit four linear models (one for quality and one for difficulty for each of the two strategies) to examine whether spatial visualization ability, Need for Cognition scores, and standardized mathematics scores predicted factor scores for quality and difficulty for the Gauss and area strategies. For both strategies, participants with higher Need for Cognition scores rated the strategies as less difficult than participants with lower Need for Cognition scores, Gauss: $t(133) = -2.75$, $p = .007$, $\eta^2 = .059$; area: $t(238) = 2.65$, $p = .008$, $\eta^2 = .029$. For the area strategy, participants with higher Need for Cognition scores also rated the strategy higher in quality, $t(238) = 2.03$, $p = .043$, $\eta^2 = .017$, and participants with higher spatial visualization abilities rated it as less difficult, $t(238) = 2.00$, $p = .047$, $\eta^2 = .016$. No other effects were significant.

3.5.2. Ratings and general strategy preferences. At the end of the study, we included three questions about participants' general strategy preferences. Specifically, we asked how much they liked to use short-cuts, strategies they understand well, and strategies that have few steps. We wanted to examine whether these general strategy preferences predicted strategy adoption, over and above the factors included in our pre-registered model, and for this reason we limited our analysis to participants who used summation at pretest (the same sample used in our pre-registered analyses). Table 5 presents the correlations between responses to these strategy preference items and the individual difference characteristics that we measured. Ratings of liking to use short-cuts and strategies with few steps were correlated. However, neither correlated with ratings of liking to use strategies that were understood well. Need for Cognition scores were negatively related to preference for shortcuts and strategies with few steps. Spatial visualization ability was positively related to preferences for strategies that were understood well. Table 6 presents correlations between these general strategy preferences and quality and difficulty ratings for the two target strategies. Overall, participants' general strategy preferences were not related to their ratings of the quality and difficulty of the strategies, with the exception that participants who rated the Gauss strategy as high in quality also tended to report liking to use strategies that they understood well.

We also explored whether these general strategy preferences predicted strategy adoption, over and above the individual difference characteristics and strategy ratings that we examined in our preregistered analyses. None of the general strategy preferences was a significant predictor. However, this lack of effect should be considered with caution, given that we measured each construct using only one item.

Table 5. Correlations between strategy preferences and learner characteristics. ** $p \leq .01$

	Shortcuts	Few steps	Understand well	Spatial visualization ability	Need for Cognition score
Few steps	.44**				
Understand well	-.14	.07			
Spatial visualization ability	-.17	-.05	.21**		
Need for Cognition score	-.46**	-.21**	.03	.36**	
Standardized math score	-.15	-.09	.00	.15	.14

Table 6. Correlations between strategy preferences and ratings of strategies. * p < .05

	Shortcuts	Few steps	Understand well
Gauss Quality	.06	.11	.22*
Gauss Difficulty	.02	-.04	.07
Area Quality	.12	.09	-.08
Area Difficulty	.09	.02	-.04

4. Discussion

Our study replicates past findings that people do not adopt all correct strategies at similar rates, and it provides evidence that this differential adoption is guided by people's evaluations of the strategies. Indeed, people's initial evaluations of the strategies predicted which strategies they adopted, over and above individual difference measures and baseline rates of adoption for each strategy. Further, ratings of the quality of the strategies were more strongly associated with strategy adoption than ratings of the difficulty of the strategies. Additionally, participants' evaluations of the strategies were not uniform, and they depended on individual difference characteristics. Participants high in need for cognition rated the quality of both strategies higher than participants low in need for cognition. Additionally, participants with higher spatial visualization ability viewed the area strategy, which involved a visual representation, as less difficult than participants with lower spatial visualization ability.

4.1 Theoretical implications

Given the importance of strategy evaluations in predicting strategy adoption, our findings suggest that models of strategy change should take people's evaluations of strategies into account. Most past models of strategy change (Lovett & Schunn, 1999; Rieskamp & Otto, 2006) do not incorporate information about the target strategies, let alone about people's evaluations of those strategies. Even when models do incorporate strategy information, often the only information included is whether a strategy is correct or incorrect (e.g., Shrager & Siegler, 1998; Siegler & Araya, 2005). Our work suggests that models of strategy change should also incorporate information about the perceived quality of the strategy. When people think a strategy is more common, more intuitive, and more efficient, they are more likely to adopt it.

Theories of strategy change should also address how contextual, individual, and strategy factors are integrated during problem solving. Our study highlights the importance of all three of these classes of factors. Our main finding was that people's evaluations of strategies were critical in determining whether they would adopt those strategies. Specifically, we found that participants were more likely to adopt strategies when they judged those strategies as higher quality. Future work should examine how people's evaluations of strategies influence not only *whether* they use the strategy (as we do here), but *when* they choose to use it.

It is worth noting that we did not assess whether the target strategies were in fact novel for all participants. We assume that the strategies were novel because relatively few participants used them at pretest, even though they are faster to implement, less error prone, and match the problem framing better than the summation strategy. However, the strategies might have not been completely novel for all participants, meaning that their prior experiences with the strategies might have influenced their adoption and possibly their ratings of these strategies. However, even if this is the case, our study still demonstrates the importance of considering strategy-level factors beyond correctness in order to understand processes of strategy use and strategy change.

We also found that features of the problem context influenced participants' strategy choices. Specifically, participants opted to use the summation strategy at posttest more frequently for the discretely-framed problem than for the continuously-framed one. This finding aligns with prior work showing that people adaptively select strategies by attending to features of the problems (Alibali et al., 1999; Lemaire & Siegler, 1995; Walsh & Anderson, 2009).

We also found that individual characteristics, including participants' confidence in their pretest strategy, need for cognition, spatial visualization ability, and general strategy preferences were associated with whether participants adopted the strategies to which they were exposed. It seems likely that people's confidence in their pretest strategy influences their willingness to abandon or let go of that strategy in favor of something new. Need for cognition may influence people's overall willingness to try to something new, and their general strategy preferences and spatial visualization ability may influence their willingness to try the specific strategy to which they were exposed. In future work, it may be valuable to try to distinguish factors that promote abandoning strategies and factors that promote adopting strategies.

We also found that certain combinations of individual characteristics were associated with strategy adoption. As in prior work, we observed a significant interaction between participants' confidence in their pretest strategies and their need for cognition on some measures of strategy adoption. However, the specifics of this interaction varied from that reported in prior work. We found that, among participants with lower need for cognition, those who had lower confidence in their pretest strategy were more likely to adopt the Gauss strategy, but among participants with higher need for cognition, their likelihood of adopting the Gauss strategy did not depend on level of confidence. This data pattern is similar to that observed for adoption of the Gauss strategy by Brown et al. (2019). However, Brown et al. (2019) also reported an interaction of confidence and need for cognition for the *area* strategy, such that participants who had both low confidence in their pretest strategy and high need for cognition were highly likely to adopt the area strategy. This pattern was not observed in the present study, perhaps because adoption of the area strategy was much less frequent than in Brown et al. (2019), presumably due to procedural differences between the studies. Notably, participants in the present study did not receive any feedback about their pretest strategies, whereas half of Brown et al.'s participants were informed that their pretest strategy was incorrect. More generally, our findings suggest that need for cognition and confidence in prior strategies may interact to influence patterns of strategy adoption, but the specifics of this interaction may depend on other factors, such as the provision of feedback about whether prior strategies are correct.

Our findings indicate that a wide range of individual difference characteristics—including not only confidence in existing strategies and need for cognition, but also mathematics ability and spatial visualization ability—are relevant to strategy selection. Although we did not find that people's general strategy preferences were related to strategy adoption, this could be due to the fact that we measured each of these preferences with only one item. It is possible that we might have observed a relation between these constructs and strategy adoption, if we had used better measures of these constructs. Additionally, other strategy preferences beyond the ones we examined might also be related to adoption. As one example, people's preference for using inventive or untaught strategies to solve problems, termed "bold problem solving" (Lubienski et al., 2021), might influence their patterns of strategy adoption. People high in bold problem solving might favor adopting strategies they see an innovative or unusual, whereas people low in bold problem solving might choose instead to adopt strategies that seem intuitive.

More generally, individual difference characteristics and features of strategies may interact to influence strategy adoption.

Our study provides evidence that contextual, individual, and strategy factors interact in important ways. For example, we found that some individual difference characteristics were related to strategy evaluations, suggesting that some individual differences may influence strategy adoption by positively or negatively influencing people's initial evaluations of novel strategies—a possibility that could be tested in future work with mediational models. Similarly, contextual factors might influence people's evaluations of strategies and thereby influence strategy adoption. For example, we found that participants were less likely to use the area strategy on a discretely-framed problem than on a continuously-framed problem. Might the area strategy be perceived as lower quality if it were presented with a discretely-framed problem than if it were presented with a continuously-framed problem? And might such differences in strategy evaluations subsequently affect strategy selection? More research is needed on how factors at multiple levels of analysis interact to predict strategy choice and change (see Alibali et al., 2019, for discussion).

4.2 Educational implications

The finding that people's initial evaluations of strategies matter for strategy adoption also has implications for educational practice. Our results suggest that simply presenting a strategy might not be enough to get students to adopt it. However, if instructors spend some time conveying to students why the strategy is of high quality (for instance, by explaining why the strategy makes sense), then students might be more likely to adopt it. This idea suggests an important direction for future research: it would be valuable to understand how changing students' perceptions of the quality of a strategy influences their strategy adoption.

Our findings are also relevant to the literature on conceptual and procedural knowledge. Research on mathematical learning holds that accurately implementing a correct strategy reflects procedural knowledge, whereas understanding *why* certain strategies work is a form of conceptual knowledge (Crooks & Alibali, 2014). Both forms of knowledge are critical for success in mathematics (Canobi, 2009; Rittle-Johnson, 2017), and there has been extensive debate regarding the proper order in which instruction should focus on these two forms of knowledge (see, e.g., Rittle-Johnson et al., 2015). Some research suggests that providing students with relevant conceptual knowledge prior to introducing problem-solving procedures leads to deeper learning (Rittle-Johnson & Alibali, 1999). The present findings suggest that one possible reason for this may be that conceptual knowledge allows students to make more informed judgments regarding the quality of new strategies, leading to better strategy choices.

4.3. Limitations

Several important limitations to this study must be acknowledged. Our sample was predominantly white and was drawn from a moderately selective university, and participants completed the study in a laboratory setting. Our results may not generalize to different samples or to actual classroom settings. Most participants in this study had previously taken at least one semester of calculus, so the algebra problems used in this study were well within their abilities, but our findings also might not generalize to samples of students with different levels of experience with mathematics. Students with more mathematics experience might be more likely to understand that the strategies lead to the same outcome and therefore judge them as similar in quality. Students with less mathematics experience might find the problems and the strategies more challenging to understand, so they might focus more on the difficulty of the strategies.

Second, our posttest was short, consisting of only two problems, and it took place immediately after the strategy exposure. It is possible that participants would have been more likely to adopt the target strategies, if they had had more chances to do so.

Third, we do not have a way of verifying that the strategies were indeed novel to our participants. However, in this regard, it is worth noting that both the Gauss and area strategies are more efficient than the summation strategy, so it seems likely that participants who were previously familiar with these strategies would have used them at pretest. On this basis, we infer that participants who used only the summation strategy at pretest had most likely not been previously exposed to the Gauss or area strategies. Furthermore, prior work shows that people frequently use the Gauss and area strategies for problems that are continuously framed, which again suggests that participants who were previously familiar with these strategies would likely have attempted them to use them at pretest. However, we did not directly measure how familiar each strategy was to participants, so we cannot be sure that the strategies were completely novel for all participants.

Finally, our study showed participants two correct and somewhat simple strategies. We did not find associations between difficulty ratings and strategy adoption, but this may be because participants perceived the strategies as similar in difficulty, and because the range of difficulty was fairly restricted. Presenting more complicated strategies or manipulating the current strategies to make them seem more difficult might provide better evidence regarding whether difficulty matters.

4.4. Conclusions

Strategy change plays an important role in cognition, development, and education. Therefore, understanding why and when people adopt new strategies for solving problems is of great importance. Research on strategy change suggests that the adoption of a new strategy depends on a myriad of contextual and individual factors. In this work, we proposed and found support for the idea that people's evaluations of the strategies themselves also play a key role in shaping patterns of strategy adoption. Specifically, people's evaluations of the *quality* of novel strategies predicted whether they adopted those strategies, over and above individual characteristics and baseline adoption rates for each strategy. Our findings suggest that classical models of strategy change that neglect strategy evaluations are missing an integral piece of the puzzle. To understand why people adopt new strategies, we need to understand how they evaluate those strategies, and why.

References

Adamson, R. E. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology*, *44*(4), 288-291. https://doi.org/10.1037/h0062487

Aguilar-Lleyda, D., Lemarchand, M., & de Gardelle, V. (2020). Confidence as a priority signal. *Psychological Science, 31*(9), 1084-1096. https://doi.org/10.1177/0956797620925039

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*(1), 127-145.

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science*, *10*(4), 327-333. https://doi.org/10.1111/1467-9280.00163

Alibali, M. W., & Booth, T. L. (2002). Representation strength influences strategy use and strategy discovery. In C.D. Schunn, & W. Gray (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 59-64). Austin, TX: Cognitive Science Society.

Alibali, M. W., Brown, S. A., & Menendez, D. (2019). Understanding strategy change: Contextual, individual, and metacognitive factors. *Advances in Child Development and Behavior, 56,* 227-256. https://doi.org/10.1016/bs.acdb.2018.11.004

Alibali, M. W., Crooks, N. M., & McNeil, N. M. (2018). Perceptual support promotes strategy generation: Evidence from equation solving. *British Journal of Developmental Psychology, 36(2)*, 153-168.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*(2), 181-214.

Barbieri, C., & Booth, J. L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences, 48*, 36-44.

Bartel, A. N., & Alibali, M. W. (2021). Symbolizing algebraic story problems: Are diagrams helpful? *Applied Cognitive Psychology, 35, 1427–1442.* https://doi.org/10.1002/acp.3874

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology*, *101*(2), 282-293. https://doi.org/10.1037/a0014312

Booth, J. L., McGinn, K. M., Young, L. K., & Barbieri, C. (2015). Simple practice doesn't always make perfect: Evidence from the worked example effect. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 24-32.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389. https://doi.org/10.1037/met0000159

Brown, S. A., & Alibali, M. W. (2018a). Promoting strategy change: Mere exposure to alternative strategies helps, but feedback can hurt. *Journal of Cognition and Development*, *19*(3), 301-324. https://doi.org/10.1080/15248372.2018.1477778

Brown, S. A., & Alibali, M. W. (2018b). What kind of problem is this? Labels guide generalization of math strategies. In C. Kalish, M. Rau, T. Rogers, and J. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (p. 2818). Austin, TX: Cognitive Science Society.

Brown, S. A., Menendez, D., & Alibali, M. W. (2018). How do people evaluate problem-solving strategies? Efficiency and intuitiveness matter. In C. Kalish, M. Rau, T. Rogers, and J. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 1414-1419). Austin, TX: Cognitive Science Society.

Brown, S. A., Menendez, D., & Alibali, M. W. (2019). Strategy adoption depends on dharacteristics of the instruction, learner, and strategy. *Cognitive Research: Principles and Implications, 4,* 1-18. https://doi.org/10.1186/s41235-019-0158-3

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306-307. https://doi.org/10.1207/s15327752jpa4803_13

Canobi, K. H. (2009). Concept–procedure interactions in children's addition and subtraction. *Journal of Experimental Child Psychology, 102(2),* 131-149. https://doi.org/10.1016/j.jecp.2008.07.008

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, *34*(4), 344-377. https://doi.org/10.1016/j.dr.2014.10.001

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, *58*(5). i-113. https://doi.org/10.1037/h0093599

Durkin, K., Rittle-Johnson, B., Star, J. R., & Loehr, A. (2021). Comparing and discussing multiple strategies: An approach to improving algebra instruction. *The Journal of Experimental Education,* 1-19.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests (revised edition)*. Educational Testing Service, Princeton, NJ.

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 1-16. https://doi.org/10.1037/xlm0000153

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*(4), 430.

Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *Elife*, *5*, e13388. https://doi.org/10.7554/eLife.13388

Gutiérrez, J. F., Brown, S. A., & Alibali, M. W. (2018). Relational equity and mathematics learning: Mutual construction during collaborative problem solving. *Journal of Numerical Cognition*, *4*(1), 159-187. https://doi.org/10.5964/jnc.v4i1.91

Hainguerlot, M., Vergnaud, J. C., & De Gardelle, V. (2018). Metacognitive ability predicts learning cue-stimulus associations in the absence of external feedback. *Scientific Reports*, *8*(1), 1-8. https://doi.org/10.1038/s41598-018-23936-9

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science, 3*(3), 446-474. https://doi.org/10.1111/j.1756-8765.2011.01150.x

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*(4), 684-689. https://doi.org/10.1037/0022-0663.91.4.684

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition*, *22*(4), 411-430. https://doi.org/10.3758/BF03200867

Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences*, *9*(1), 19-42. https://doi.org/10.1016/S1041-6080(97)90018-2

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children’s learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83–97.

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology: General*, *128*(2), 107-130. https://doi.org/10.1037/0096-3445.128.2.107

Lubienski, S. T., Ganley, C. M., Makowski, M. B., Miller, E. K., & Timmer, J. D. (2021). “Bold problem solving”: A new construct for understanding gender differences in mathematics. *Journal for Research in Mathematics Education, 52*(1), 12-61.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, (6, No. 248).

McNeil, N. M. (2014). A change–resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives, 8*(1), 42-47.

Newcombe, N.S. (2018). Three kinds of spatial cognition. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed.) (pp. 521–552). Wiley.

Newcombe, N. S., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying visual and spatial reasoning for design creativity* (pp. 179–192). New York, NY: Springer. https://doi.org/10.1007/978-94-017-9297-4_10.

Paas, F., & van Merriënboer J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*, 122-133.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207-236. https://doi.org/10.1037/0096-3445.135.2.207

Riggs, A. E., Alibali, M. W., & Kalish, C. W. (2015). Leave her out of it: Person-presentation of strategies is harmful for transfer. *Cognitive Science, 39*(8), 1965-1978. https://doi.org/10.1111/cogs.12224

Riggs, A. E., Alibali, M. W., & Kalish, C. W. (2017). Does it matter how Molly does it? Person-presentation of strategies and transfer in mathematics. *Contemporary Educational Psychology*, *51*, 315-320. https://doi.org/10.1016/j.cedpsych.2017.09.001

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives, 11(3),* 184-190. https://doi.org/10.1111/cdep.12229

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91(1),* 175. https://doi.org/10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, *27*(4), 587-597. https://doi.org/10.1007/s10648-015-9302-x

Sadowski, C. J., & Gülgöz, S. (1992). Internal consistency and test-retest reliability of the Need for Cognition Scale. *Perceptual and Motor Skills*, *74*(2), 610-610. https://doi.org/10.2466/pms.1992.74.2.610

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*(5), 405-410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (2000). The rebirth of children's learning. *Child Development*, *71*(1), 26-35. https://doi.org/10.1111/1467-8624.00115

Siegler, R., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior*, 33, 1-42. https://doi.org/10.1016/S0065-2407(05)80003-5

Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology*, *27*(2), 194-226. https://doi.org/10.1006/cogp.1994.1016

Siegler, R., & Jenkins, E. A. (2014). *How children discover new strategies*. Psychology Press.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, *126*(1), 71-92. https://doi.org/10.1037/0096-3445.126.1.71

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon and G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.

Star, J. R., & Rittle-Johnson, B. (2008). Flexibility in problem solving: The case of equation solving. *Learning and Instruction, 18*, 565 - 579. doi:10.1016/j.learninstruc.2007.09.018

Stieff, M., Ryu, M., Dixon, B., & Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. *Journal of Chemical Education, 89*(7), 854-859.

Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics? *Current Directions in Psychological Science*, *22*(5), 367-373. https://doi.org/10.1177/0963721413484756

van der Ven, S. H., Boom, J., Kroesbergen, E. H., & Leseman, P. P. (2012). Microgenetic patterns of children's multiplication learning: Confirming the overlapping waves model by latent growth modeling. *Journal of Experimental Child Psychology*, *113*(1), 1-19. https://doi.org/10.1016/j.jecp.2012.02.001

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817-835. https://doi.org/10.1037/a0016127

Walsh, M. M., & Anderson, J. R. (2009). The strategic nature of changing your mind. *Cognitive Psychology*, *58*(3), 416-440. https://doi.org/10.1016/j.cogpsych.2008.09.003