



Published in final edited form as:

Cogn Dev. 2022 ; 62: . doi:10.1016/j.cogdev.2022.101167.

Children's confidence using incorrect strategies on mathematical equivalence problems

Amanda Grenell,
Indiana University

Lindsey J. Nelson,
Indiana University

Bailey Gardner,
Indiana University

Emily R. Fyfe
Indiana University

Abstract

Children often struggle to solve mathematical equivalence problems correctly. The change-resistance theory offers an explanation for children's difficulties and suggests that some incorrect strategies represent the overgeneralization of children's narrow arithmetic experience. The current research considered children's metacognitive abilities to test a tacit assumption of the change-resistance theory by providing a novel empirical examination of children's strategy use and certainty ratings. Children were recruited from U.S. elementary school classrooms serving predominantly White students between the ages of 6 and 9. In Study 1 ($n = 52$) and Study 2 ($n = 147$), children were more certain that they were correct when they employed arithmetic-specific incorrect strategies relative to other incorrect strategies. These findings are consistent with the change-resistance theory and have implications for the development of children's metacognition.

Keywords

Metacognition; Mathematical Equivalence; Strategy Use; Early Childhood

Address correspondence to Amanda Grenell, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th Street, Bloomington IN 47405. agrenell@iu.edu.

Author Note: Amanda Grenell, Lindsey Nelson, Bailey Gardner, and Emily R. Fyfe, Department of Psychological and Brain Sciences, Indiana University.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interest: None

¹The two sets of problems were well-matched. Children's overall accuracy for the problems in Set 1 and Set 2 did not significantly differ after controlling for their prior knowledge about mathematical equivalence, which was measured as part of the larger study. More information about the criteria used to select and match the items in Sets 1 and 2 and children's accuracy on the items in Set 1 versus Set 2 can be found in Nelson & Fyfe (2019).

Children often use a variety of different strategies to solve a target problem (e.g., Alibali et al., 2019; Fazio et al., 2016; Lemaire, 2017), and these strategies often vary in key ways. Some strategies produce a correct answer and others produce an incorrect answer, some are used frequently and some are used sparingly. These variations provide insights into children's developing knowledge. Differences in strategy use have led to advancements in theories on cognitive development (e.g., Siegler, 2000), the characterization of learning disabilities (e.g., Geary et al., 2007), and the design of teaching practices (e.g., Rittle-Johnson et al., 2020).

One area of research that has been influenced by variations in strategy use is children's understanding of mathematical equivalence problems (e.g., Alibali, 1999; Hornburg et al., 2018; Matthews & Rittle-Johnson, 2009). These problems often contain operations on both sides of the equal sign (e.g., $3 + 4 = 5 + \underline{\quad}$) and tap into children's early algebraic reasoning (e.g., Kieran, 1981). Children's performance on mathematical equivalence problems has led to the development of a "change-resistance" account that helps explain why some strategies become entrenched and difficult to let go (McNeil, 2014). The current study aims to shed light on this theoretical account by providing a novel empirical examination of children's strategy use and metacognitive abilities on mathematical equivalence problems.

Metacognitive Abilities and Mathematical Problem Solving

Metacognition is the ability to think about one's own knowledge (Nelson & Narens, 1990), and it includes the ability to monitor and evaluate one's performance (e.g., "Did I solve that problem correctly?" "How sure am I about my answer?"). Metacognitive skills are positively associated with children's academic achievement, including in mathematics (e.g., Nelson & Fyfe, 2019; Bellon et al., 2019; Vo et al., 2014). Previous research suggests that even young children are capable of monitoring their knowledge and differentiating between answers that are correct versus incorrect (e.g., Coughlin et al., 2015; Desoete et al., 2006; Vo et al., 2014).

Previous research on children's metacognitive skills tends to focus solely on correct versus incorrect answers. However, it is possible that children's metacognitive monitoring may differ depending on the specific type of strategy they use. Monitoring different *incorrect* strategies may be particularly challenging as children tend to overestimate their knowledge, often thinking they are correct when they are not (e.g., Finn & Metcalfe, 2014; Flavell, 1979; Rinne & Mazzocco, 2014; Roebbers & Spiess, 2017). While overconfidence may have some benefits for learning, such as allowing children to be open to learning new and challenging topics (e.g., Bjorklund & Green, 1992), overconfidence is also negatively related to children's metacognitive abilities (e.g., Destan & Roebbers, 2015). Monitoring different incorrect strategies may also require introspection about multiple cognitive states. For example, some incorrect strategies stem from a complete *lack* of knowledge (e.g., guessing), while other incorrect strategies stem from the *presence* of inaccurate knowledge (e.g., misapplying a strategy that works in another context). This suggests that incorrect strategies are not all cognitively equivalent and that an examination of children's metacognitive monitoring abilities can be a fruitful way to shed light on children's strategy knowledge (e.g., Durkin & Rittle-Johnson, 2015), particularly in areas where misconceptions are prevalent, such as in the domain of mathematical equivalence.

Mathematical Equivalence Problems

Mathematical equivalence is the idea that two quantities are the same and it is often formally represented by the equal sign (e.g., Kieran, 1981). Knowledge of mathematical equivalence is considered foundational for both arithmetic and algebra (e.g., Carpenter et al., 2003; Knuth et al., 2006). It enables children to focus on the *relation* between quantities and to understand that the two sides of an equation represent the same amount (e.g., $3 + 4 = 7$; $3 + 4 = 5 + 2$). Children's knowledge of mathematical equivalence predicts their mathematics achievement and algebra competence in later grades (Matthews & Fuchs, 2020; McNeil et al., 2019).

One way children demonstrate their knowledge of equivalence is by using a correct strategy to solve *mathematical equivalence problems* (e.g., Perry et al., 1988). In line with Rittle-Johnson et al. (2011), we use the term “mathematical equivalence problems” to refer to a broad set of problems that include operations on the right side of the equal sign (e.g., $8 = 6 + \underline{\quad}$) or operations on both sides of the equal sign (e.g., $3 + 4 = 5 + \underline{\quad}$). Critically, these problems are distinct from traditional arithmetic problems with operations on the left side only (e.g., $6 + 2 = \underline{\quad}$). Unfortunately, correct strategies on these type of mathematical equivalence problems are relatively rare. For children ages 7 to 11 in the United States, success rates on mathematical equivalence problems are modest at best, with rates as low as 15–25% on pretest measures that occur prior to any training or instruction (e.g., Fyfe & Rittle-Johnson, 2017; McNeil & Alibali, 2005; Rittle-Johnson, 2006). For that reason, most research focuses on children's difficulties with mathematical equivalence and the *incorrect* strategies that children use.

The Change-Resistance Account

A change-resistance theory has emerged as a widely-accepted explanation as to why children exhibit difficulties with mathematical equivalence (McNeil, 2014). The theory suggests that children's difficulties arise largely because of their prior experiences with arithmetic (Knuth et al., 2006; McNeil & Alibali, 2005). The idea is that children pick up on common patterns in their arithmetic practice and develop efficient strategies that are consistent with those patterns. These strategies can become entrenched and then misapplied to new problems that share some, but not all, characteristics with their arithmetic practice problems.

In the United States, children are often exposed to narrow arithmetic practice in an “operations-equal-answer” format (e.g., $2 + 3 = 5$; McNeil et al., 2006). One analysis of mathematics textbooks found that these traditional arithmetic formats typically accounted for over 90% of all equations across grades 1 to 5 (Powell, 2012). Given this exposure, children learn to expect certain patterns – the operations will always be on the left side and the equal sign is followed by the answer. Consistent with the change-resistance theory, many children view the equal sign as a signal to “add the numbers” as opposed to a symbol relating two amounts (e.g., Baroody & Ginsburg, 1983; Behr et al., 1980; Kieran, 1981; Stephens et al., 2013), and they develop arithmetic-specific strategies that they

overgeneralize when they encounter novel mathematical equivalence problems (e.g., McNeil & Alibali, 2005; Rittle-Johnson, 2006).

Prior work has documented children's use of these incorrect strategies (Perry et al., 1988). Consider the problem $3 + 4 = 5 + \underline{\quad}$. The first arithmetic-specific strategy is *add-all* in which children add all the numbers in the problem (e.g., 3 plus 4 plus 5 adds to 12). The second arithmetic-specific strategy is *add-to-equal* in which children add all the numbers before the equal sign (e.g., I saw the 3 and 4 and I wrote 7 in the blank). These two strategies work well on traditional arithmetic problems but produce incorrect solutions on equivalence problems. Children also use other incorrect strategies to solve equivalence problems. For example, some children use a *carry* strategy by copying a number from the problem (e.g., I saw 3 here, so I put 3 in the blank; Alibali, 1999; Fyfe et al., 2012; McNeil & Alibali, 2005). However, consistent with the change-resistance account, the two arithmetic-specific strategies are the most prevalent and often account for 60–90% of children's incorrect responses (e.g., Matthews & Rittle-Johnson, 2009; McNeil & Alibali, 2005; Rittle-Johnson, 2006). When children encounter mathematical equivalence problems, they see features that activate their arithmetic knowledge (e.g., addends, equal sign), and they overgeneralize strategies that have become entrenched through their narrow experience with traditional arithmetic.

The Role of Children's Metacognitive Abilities

One tacit assumption of the change-resistance account is that children are confident that the arithmetic-based strategies are a correct approach. That is, the misapplication of these arithmetic-specific strategies to mathematical equivalence problems is not an uncertain, in-the-moment guess, but a firm and confident reliance on a familiar strategy that has worked well in the past to solve traditional arithmetic problems. In the now seminal paper on the change-resistance account by McNeil and Alibali (2005), the authors use confidence as an indicator of entrenchment. Specifically, they had children solve three mathematical equivalence problems and rate "how certain they were about their way of doing the problem" using a 1 to 7 scale after each one (McNeil & Alibali, 2005, p.886). To quantify entrenchment, the child had to use the incorrect *add-all* strategy on at least 2 of the 3 problems *and* report high confidence in the correctness of that strategy (i.e., rate greater than 4 on a 1–7 scale). Children who met these criteria were said to be operating according to an entrenched pattern instead of a generic fall-back rule – a strategy that children apply across a wide variety of problems when they are not sure what to do (Siegler, 1983). Thus, these authors identify frequency of use and certainty in correctness as indicators of entrenchment based on the change-resistance account.

One untested prediction that stems from this assumption is that children should be more certain when using arithmetic-specific strategies relative to other incorrect strategies. According to the change-resistance account, arithmetic-specific strategies are likely to stem from the overgeneralization of entrenched, familiar patterns that children encounter in arithmetic, whereas other incorrect strategies are likely to stem from in-the-moment processing of item-level information (e.g., copying a number they see). Further, McNeil and Alibali (2005) propose that high confidence is associated with using entrenched strategies.

In contrast, if the arithmetic-specific strategies are not cognitively unique and simply represent another way to solve the problems incorrectly, then children's certainty should not be systematically higher on these arithmetic-specific strategies relative to other strategies.

Of course, there are other features of these strategies that might influence children's certainty. For example, typically the more time people spend implementing a strategy the lower their certainty is in the correctness of their response (e.g., Coughlin et al., 2015). The idea is that people use their response time as a cue to their lack of knowledge or to the amount of doubt they have on a given task (e.g., Koriat et al., 2006). Thus, strategies that take longer to execute on mathematical equivalence problems may result in lower certainty. Critically, there are no systematic reasons to think that the arithmetic-specific strategies are faster to implement relative to other incorrect strategies. In fact, a child who adds all the numbers in the problem (e.g., answering 12 for $3 + 4 = 5 + \underline{\quad}$) may take longer than a child who simply copies a number they see (e.g., answer 3 for $3 + 4 = 5 + \underline{\quad}$), in which case they may be less certain using the *add-all* strategy. Thus, apart from the change-resistance account, there is no systematic reason to expect children's certainty to be markedly higher on the *add-all* and *add-to-equal* strategies.

The Current Study

The goal of the current study is to test a prediction of the change-resistance account by providing a novel empirical examination of children's strategy use and metacognitive abilities (i.e., certainty ratings) on mathematical equivalence problems. There are three published studies that include a measure of certainty rating on equivalence problems, but none of them focus on the match between certainty and different strategy types. As mentioned above, McNeil and Alibali (2005) quantified children's adherence to arithmetic patterns, which included (among other things) using the *add-all* strategy on pretest problems and giving it a certainty rating of 5 or higher on a 7-point scale. They found that higher adherence to arithmetic patterns was negatively associated with learning from a lesson on equivalence. Brown and Alibali (2018) had children rate their certainty using a 5-point scale on a mathematical equivalence problem prior to a brief lesson. For children with lower certainty, feedback increased the likelihood that children would change their strategy after the lesson. Nelson and Fyfe (2019) examined children's monitoring and help-seeking on equivalence problems and their associations with a broader measure of mathematics knowledge. To calculate a measure of monitoring, children rated their confidence on a 4-point scale for five mathematical equivalence problems. Children reported a confidence level consistent with the correctness of their response only 57% of the time, but individual differences in this type of monitoring was positively related to their mathematics knowledge.

The current research includes two studies in which elementary school children solve a set of mathematical equivalence problems and rate their certainty on each problem. The goal is to examine children's certainty as a function of strategy use to test an assumption of the change-resistance account. Study 1 is a re-analysis of a portion of the data presented in Nelson and Fyfe (2019) with a novel focus on specific strategy types. Study 2 is a re-analysis of data from an independent, larger sample of elementary school children to serve as a replication of the findings from the first study (Fyfe et al., 2021).

The first prediction is that the data will replicate prior findings (e.g., McNeil & Alibali, 2005; Rittle-Johnson, 2006) and demonstrate that the two arithmetic-specific strategies (i.e., *add-all* and *add-to-equal*) are the most common incorrect strategies that children use on equivalence problems. The second prediction is that children's certainty will be higher on the arithmetic-specific strategies relative to other incorrect strategies because the arithmetic-specific strategies stem from entrenched, routine patterns encountered in children's mathematics experience. The results will shed light on how variations in children's strategy use can provide insights into their cognitive development and their difficulties with mathematical equivalence problems.

Study 1

Method

Participants—Children were recruited from four elementary schools in a U.S. Midwest college town, including two private schools, one Montessori school, and one public charter school. Parent consent was obtained from 54 children across five participating classrooms (see Supplementary Material for a justification of the sample size). Two children failed to complete the study, so the final analytic sample included 52 elementary school children with an average age of 8;1 years ($SD = 0.8$). This included 14 children in first grade ($M = 7;2$ years, $SD = 0.4$), 22 children in second grade ($M = 8;0$ years, $SD = 0.4$), and 16 children in third grade ($M = 9;1$ years, $SD = 0.4$). On the consent form, parents had the option of providing information about their child's gender and ethnicity through open-ended fill-in-the-blank questions. Thirty-eight parents opted to provide gender information by writing female, male, F, or M; 17 of the 38 were female. Thirty-two parents opted to provide their child's ethnicity information; 24 were White, 1 was Hispanic, 1 was Asian, and 6 were Multiracial. The same sample was reported in Nelson and Fyfe (2019) to answer a separate research question. The larger published report examined children's metacognitive skills and their association with a comprehensive measure of mathematical equivalence knowledge. It did not include an examination of children's strategy use. Thus, the hypotheses, analyses, and conclusions in the current paper are distinct from those reported in Nelson and Fyfe (2019).

Materials—Children solved a set of five mathematical equivalence problems. One problem was a simpler item with an operation on the right side of the equal sign (e.g., $7 = 4 + \underline{\quad}$) and the remaining four items had operations on both sides of the equal sign (e.g., $3 + 7 = 3 + \underline{\quad}$). Two sets of isomorphic problems were constructed to ensure the results did not depend on the specific numbers used, and children were randomly assigned to solve the problems in Set 1 or Set 2¹. See Table 1 for the list of the mathematical equivalence problems. The problems were presented in a horizontal format with equal spacing between the addends, operations, and equal sign.

Children also rated their certainty immediately after solving each problem on a trial-by-trial basis using a traffic light rating scale adapted from prior research (De Clercq et al., 2000). There were four options with anchors from "I know I got this problem right" to "I know I got this problem wrong." See Figure 1 for a depiction of the scale.

Coding—Children’s certainty ratings were assigned a score of 1 to 4 with 1 corresponding to “I know I got this problem wrong” and 4 corresponding to “I know I got this problem right.” Thus, higher scores are equated with higher certainty in correctness.

Children’s strategies on the mathematical equivalence problems were based on their written solutions and coded using an established system from previous research (e.g., McNeil & Alibali, 2005; Perry et al., 1988; Rittle-Johnson, 2006). See Table 2 for examples of each strategy. Solutions were coded as reflecting a correct strategy if they made both sides of the equal sign the same amount. As in prior work, solutions within ± 1 of the correct answer were coded as a correct strategy to account for minor arithmetic mistakes (e.g., McNeil, 2007; Perry, 1991). Although different types of correct strategies have been identified in previous research (e.g., Perry et al., 1988), distinguishing between correct strategies is not possible when only relying on children’s numerical solutions because different correct strategies could be used to arrive at the same correct numerical response. Therefore, in the current study, every correct numerical solution (± 1) received the correct strategy code. Incorrect strategies could be further classified as arithmetic or other given these strategies lead to different numerical responses. Arithmetic strategies included adding all the numbers in the problem (*add-all*) or adding all the numbers before the equal sign (*add-to-equal*). Again, solutions within ± 1 of the solution that would be obtained with either arithmetic strategy were coded as reflecting that strategy (e.g., on $3 + 7 + 8 = _ + 8$, a solution of 25 would be coded as *add-all* even though the actual total is 26). Other incorrect strategies included copying a number from the problem into the blank (*carry*) or idiosyncratic responses like operating on random numbers (*idiosyncratic*). The *idiosyncratic* strategy code was used for any numerical solution that could not be derived using a *correct* strategy, the *add-all* strategy, the *add-to-equal* strategy, or the *carry* strategy. One child left one item blank; this trial was excluded from analysis as it was unclear whether it represented a deliberate choice (e.g., I don’t know how to answer that question) or a logistic mistake (e.g., accidentally skipping the problem). A second rater coded the strategies for 31% of children’s solutions and inter-rater agreement was high ($\kappa = .98$).

Given the central importance of identifying children’s different strategies for our research questions, we verified the reliability of our strategy coding in three additional ways. We summarize these ways here and provide further details in the Supplementary Materials.

We first verified that using written solutions alone is a reliable indicator of children’s strategies. Some prior work has relied solely on written solutions to infer strategy use (e.g., McNeil, 2007), especially when assessments are administered as pretests or posttests to a whole group of students (e.g., Fyfe et al., 2012; Matthews & Rittle-Johnson, 2009). But, in one-on-one contexts, it is common to rely on written solutions in conjunction with children’s verbal strategy reports of how they arrived at their solutions, and these verbal reports can provide unique information (e.g., McGilly & Siegler, 1990). Children in the current sample did not provide verbal strategy reports for the target items reported here. However, as part of the larger study, children completed a separate, more comprehensive measure of mathematical equivalence knowledge, which included a set of problems on which they provided numerical solutions and verbal strategy reports (see Nelson & Fyfe, 2019). The first author coded children’s strategy use on these items using their written solutions alone.

A second researcher independently coded children's strategy use on these items using both written solutions and verbal strategy reports. The inter-rater agreement was high ($\kappa = .91$), indicating that the written numerical solutions alone were a reliable indicator of children's strategy use for this group of children.

Additionally, we verified that the classification of children's strategies was reliable with minor changes to the coding criteria. The current coding criteria were somewhat lenient (e.g., answers within ± 1 of the *add-all* solution were coded as reflecting the *add-all* strategy). Therefore, we also coded the strategies in two complementary ways using more stringent criteria. We found that the distribution of children's strategy use was similar using these different criteria and produced similar conclusions as the original coding criteria (see Supplementary Materials).

Procedure—Children participated in a one-on-one session with a researcher in a quiet place in their school for 30 minutes. The researcher presented all problems on a computer screen using a Qualtrics survey. Children were first presented with a warm-up problem (Which number is bigger? 29 or 31) to ensure they could use the computer controls correctly and to ensure they understood how to use the certainty rating scale. Then, children solved the five mathematical equivalence problems one at a time. For each problem, they typed their numerical response and then clicked on a certainty rating to indicate how certain they were in their answer. This was repeated for all five problems, and no feedback was provided. The computer recorded children's response time on each item, which was the time between when the problem first appeared on the screen and when the child submitted their typed numerical response.

Results

Descriptive Statistics—Table 3 displays the frequency of children's certainty ratings, and Table 4 displays the frequency of each strategy type. On average, children used a correct strategy on 47% of the problems. Consistent with the hypothesis, the most common incorrect strategies were arithmetic-specific; *add-all* was used on 24% of problems and *add-to-equal* was used on 15% of problems. Children's strategy use was generally variable, with children using an average of 1.96 ($SD = .91$) different strategies (out of 5). Additionally, a subset of children showed no variability in their strategy use: 16 children used a correct strategy to solve all the problems, and 11 children used the same type of incorrect strategy to solve all the problems. Of the children that used the same type of incorrect strategy across items, three children used the *add-all* strategy only, one child used the *carry* strategy only, and seven children used arithmetic-based incorrect strategies only (*add-all* and *add-to-equal*). Children were very confident in their responses; they gave a rating of 3 (out of 4) on 43% of problems and a rating of 4 (out of 4) on 41% of problems. Despite most children using the upper end of the scale, the full range of the certainty scale was used (see Table 3).

Consistent with the hypothesis, certainty ratings descriptively depended on the strategy children used. Across all problems and all children, certainty ratings were highest when children used a Correct strategy. However, among the incorrect strategies, certainty ratings

were higher on the two arithmetic-specific strategies (*add-all* and *add-to-equal*) relative to other incorrect strategies (see Table 4).

Certainty Ratings Across Different Strategy Types—To formally examine children’s certainty ratings as a function of strategy type, we conducted mixed-effects regression analyses using the lme4 package (Bates et al., 2015) in R v. 4.0.2 (R Core Team, 2020). These analyses included strategy type (Correct, Arithmetic, Other-Incorrect) as a predictor of children’s certainty ratings. Given that items were nested within children and children gave certainty ratings across items, we first calculated the intraclass correlations (*ICCs*) to determine how much variation in certainty ratings was due to participants and items. These analyses revealed that 38.1% of the variation in certainty ratings was due to participant differences and 2.3% of the variation in certainty ratings was due to item differences. We included a random intercept for Participant and a random intercept for Item in the models. Random slopes were not included in these models. Additionally, within these models, isomorphic items in Set 1 and Set 2 were coded with the same ID (e.g., a problem in Set 1 and the corresponding isomorphic problem in Set 2 were given the same Item ID number). We report the results from these models with Satterthwaite’s method for estimating degrees of freedom in the lmerTest package (Kuznetsova et al., 2017).

Given we were interested in comparing children’s confidence when using Arithmetic strategies relative to Correct strategies and when using Arithmetic strategies relative to Other-Incorrect strategies, we examined the dummy-coded contrasts of strategy type with Arithmetic strategies as the reference group in the regression analyses. The first contrast sheds light on children’s metacognition more broadly; if children have well-calibrated metacognitive abilities, they should be more certain when using correct strategies relative to incorrect strategies. The second contrast sheds light on the change-resistance account; if the arithmetic-specific strategies stem from the overgeneralization of entrenched, familiar patterns, children should be more certain when using arithmetic-specific strategies relative to other incorrect strategies. Missing data was rare (0.4%), so we removed the missing observations from the dataset for the following analyses.

In addition to our primary model (i.e., Model 1a), we conducted three additional regression analyses. Model 1b was identical to Model 1a but controlled for response time to determine how response time related to children’s certainty ratings and whether it influenced the role of strategy type. Models 2a and 2b mirrored the first two models but were based on a subset of the items to determine if the inclusion of easier items influenced the main pattern of results. As mentioned before, children solved a set of five items: four with operations on both sides of the equal sign (e.g., $9 + 4 = 3 + \underline{\quad}$), and one easier item with operations on the right side only (e.g., $8 = 3 + \underline{\quad}$). Thus, Model 2a was identical to Model 1a, but with the easier item removed. Model 2b was identical to Model 2a, but it controlled for response time.

Models 1a and 1b: Certainty Ratings Across All Items—As can be seen in Table 5, the first contrast in Model 1a revealed no significant difference in children’s certainty ratings when using Arithmetic strategies than when using Correct strategies, $B(SE) = .03(.11)$, $p = .79$. However, the second contrast revealed that children were significantly

more certain when using Arithmetic strategies than when using Other-Incorrect strategies, $B(SE) = -.53(.14)$, $p < .001$. Thus, children's confidence depended on the specific type of *incorrect* strategy they used, supporting our hypothesis.

These results are consistent with the change-resistance theory. However, an alternative explanation for children's high confidence when using arithmetic-specific strategies is that the arithmetic-specific strategies take less time to execute than other strategies and inflate children's certainty in their correctness. Indeed, in the full sample, the average time it took children to solve a problem was significantly, negatively correlated with their average certainty, $r = -.29$, $p = .04$. To rule out this possibility, differences in response time were examined as a function of strategy type, and Model 1b was conducted controlling for response time in the analyses.

As shown in Table 4, Arithmetic strategies were not more efficient than the other strategies. In fact, children tended to take less time executing correct strategies than incorrect strategies, and both types of incorrect strategies (Arithmetic and Other) took close to one full minute per problem. Model 1b included response time as a predictor of children's certainty ratings (see Table 5). Response time was a significant predictor of children's certainty ratings, $B(SE) = -.004 (0.001)$, $p = .005$, meaning the faster they solved the item the higher certainty they reported. Consistent with Model 1a, the contrast between Arithmetic strategies and Other-Incorrect strategies remained statistically significant, $B(SE) = -.49(.14)$, $p < .001$, after controlling for response time. Thus, on average, children were more confident when using Arithmetic strategies than when using Other-Incorrect strategies after accounting for the time it took to complete the problems.

Models 2a and 2b: Certainty Ratings with Easy Item Removed—Models 2a and 2b examined whether the main pattern of results held when removing the easy item (i.e., problem with operations on the right side only). When the data from the remaining four items were analyzed, the conclusions remained the same (see Table 5). In Model 2a, the only significant contrast was between Arithmetic and Other-Incorrect strategies, $B(SE) = -.54(.15)$, $p < .001$. Children were more certain when using arithmetic-specific strategies relative to other types of incorrect strategies even after the easy item was removed. Further, Model 2b revealed Response Time was no longer a significant predictor of children's confidence, $B(SE) = -0.002 (0.002)$, $p = 0.14$. However, children were still more certain when using Arithmetic strategies than Other-Incorrect strategies on this subset of items after controlling for response time, $B(SE) = -.51(.15)$, $p < .001$. Therefore, including an easier item did not seem to be driving the main results found when analyzing all items.

Study 1 Discussion

Study 1 results were consistent with the hypotheses. Children frequently solved mathematical equivalence problems using incorrect strategies, and the most common incorrect strategies were arithmetic-specific (adding all the numbers together or adding all the numbers before the equal sign). Further, children's certainty depended on the specific type of incorrect strategy they used; children were more certain that they were correct when they employed arithmetic-specific strategies relative to other incorrect strategies. These

arithmetic-specific strategies did not take less time compared to the other strategies, and accounting for response time did not alter the conclusions. These findings are consistent with the change-resistance theory and suggest that arithmetic-specific strategies are not an uncertain, in-the-moment guess, but a confident reliance on a familiar strategy that has worked well in the past.

The results from Study 1 are promising, but they rely on a small sample of children. The goal of Study 2 was to replicate the main findings from Study 1 using a larger sample.

Study 2

Method

Participants—Children were recruited from three elementary schools in a U.S. Midwest college town, including one private school and two public schools. Opt-out consent forms were distributed to 150 children across eight participating classrooms. Two parents opted their children out of participation and an additional child was excluded from analysis for not completing the study. The final sample reported here consists of 147 children (see Supplementary Material for a justification of the sample size). Demographic data for individual participants was not collected due to the opt-out consent protocol. Because we did not have birth date information for each child, we cannot calculate their precise ages, but in the U.S., children in first grade are typically 6–7 years old and children in second grade are typically 7–8 years old. Moreover, publicly available information for the three schools indicate that the student populations ranged from 60.2–75.6% White, 7.0–8.2% Multiracial, 3.0–7.2% Black, 3.7–18.5% Asian, and 3.7–5.9% Hispanic. Additionally, classroom teachers were asked to provide class-level gender data (i.e., what percentage of the students in your class are female). The participating classrooms ranged from 32–64% female. The data reported here come from a larger study on children’s metacognitive abilities in elementary school (Fyfe et al., 2021).

Materials—Children solved a set of six mathematical equivalence problems. Two problems were simpler items with an operation on the right side of the equal sign (e.g., $7 = 4 + \underline{\quad}$) and the remaining four items had operations on both sides of the equal sign (e.g., $2 + 7 = 6 + \underline{\quad}$). See Table 1 for the list of the mathematical equivalence problems. The problems were presented in a horizontal format with equal spacing between the addends, operations, and equal sign. Children also rated their certainty on each problem using the same rating scale from Study 1. The six mathematical equivalence problems were embedded in a larger mathematics assessment containing 12 other mathematics items.

Coding—Children’s strategies on the mathematical equivalence problems were based on their written solutions and coded using the same system as in Study 1 (see Table 2). Children did not provide a solution on 4.6% of all trials. As in Study 1, these “blank” solutions were excluded from analysis as it was unclear whether they represented a deliberate choice or a logistic mistake. A second rater coded the strategies for 31% of children’s solutions and agreement was high ($\kappa = .99$). Certainty ratings were assigned scores from 1 to 4 as in Study 1.

Procedure—Children completed a paper-and-pencil assessment in their classroom during their normal mathematics class. The researcher led the students through the instructions and discussed an example item (e.g., $1 + _ = 2$) to ensure children knew how to complete the assessment packet and how to use the certainty rating scale. Then, students completed the problems and circled their certainty ratings at their own pace at their desk. No feedback was provided. Given the nature of this procedure (whole group in the classroom), no response time data was collected.

Results

Descriptive Statistics—Table 3 displays the frequency of children’s certainty ratings, and Table 4 displays the frequency of each strategy type. On average, children used a correct strategy on 58% of the problems. Similar to Study 1 and consistent with our hypothesis, the most common incorrect strategy was the *add-all* strategy, which was used on 15% of problems. The *add-to-equal* strategy was also fairly common and used on 9% of all problems. Children’s strategy use was generally variable, with children using an average of 2.16 ($SD = .94$) different strategies (out of 5). Additionally, a subset of children showed no variability in their strategy use: 42 children only used a correct strategy, and one child used the same incorrect strategy (*Idiosyncratic*) to solve all the problems. Again, children were very confident in their responses; they gave a rating of 3 (out of 4) on 25% of problems and a rating of 4 (out of 4) on 54% of problems. Despite most children using the upper end of the scale, the full range of the certainty scale was used (see Table 3). Similar to Study 1, certainty ratings descriptively depended on the strategy children used. Across all problems and all children, certainty ratings were highest when children used a Correct strategy. However, among the incorrect strategies, certainty ratings were higher on the two arithmetic-specific strategies (*add-all* and *add-to-equal*) relative to other incorrect strategies (see Table 4).

Certainty Ratings Across Different Strategy Types—The same analytic plan as Study 1 was used, except the analysis controlling for response time was not included since these data were not available for Study 2. This means children’s certainty ratings were examined using two mixed-effects regression models, one with all the items included (Model 1) and one with the simpler items removed (Model 2). Additionally, we calculated the *ICCs* to determine how much variation in certainty ratings was due to participants and items. These analyses revealed that 33.9% of the variation in certainty ratings was due to participant differences and 11% of the variation in certainty ratings was due to item differences. We included a random intercept for Participant and a random intercept for Item in the models. Random slopes were not included in these models. The amount of missing data for Study 2 was fairly rare (5.1%), so missing observations were removed from the dataset when running the following analyses.

Model 1: Certainty Ratings Across All Items—As can be seen in Table 6, the first contrast in Model 1 revealed that children were significantly more certain when they used a Correct strategy than when they used Arithmetic strategies, $B(SE) = .44(.07)$, $p < .001$. The second contrast revealed that children were significantly more certain when they used Arithmetic strategies than when they used Other-Incorrect strategies, $B(SE) = -.35(.09)$, $p <$

.001. Thus, consistent with Study 1 and with our hypothesis, children's confidence depended on the specific type of incorrect strategy they used.

Model 2: Certainty Ratings with Easy Items Removed—Model 2 examined whether the main pattern of results held when removing the two easy items (i.e., problems with operations on the right side only). When the data from the remaining four items were analyzed, the conclusions remained the same (see Table 6). The first contrast showed that children were still more certain when they used Correct strategies relative to Arithmetic strategies on this subset of items, $B(SE) = .34(.09)$, $p < .001$. Similarly, the second contrast showed that children were also still more certain when they used Arithmetic strategies than when they used Other-Incorrect strategies, $B(SE) = -.39(.10)$, $p < .001$. Therefore, consistent with the findings from Study 1, including easier items did not seem to be driving the main pattern of results found when analyzing all the items.

Study 2 Discussion

Study 2 results were consistent with the hypotheses. Children frequently solved mathematical equivalence problems using incorrect strategies, and common incorrect strategies were arithmetic-specific (adding all the numbers together or adding all the numbers before the equal sign). Further, children were more certain that they were correct when they employed arithmetic-specific strategies relative to other incorrect strategies. Thus, these findings are consistent with the change-resistance theory.

General Discussion

The current study represents a novel empirical examination of children's strategy use and certainty ratings on mathematical equivalence problems that provides support for the change-resistance theory. In two independent samples, children were generally more confident when using correct than incorrect strategies (though this only reached statistical significance in Study 2). More robustly, children's common incorrect strategies were *add-all* and *add-to-equal*, which represent the misapplication of strategies that work well on traditional arithmetic problems. Further, children reported significantly higher certainty ratings when they used arithmetic-specific strategies relative to other incorrect strategies. These results were reliable across variations in analyses and in strategy-coding criteria (See Supplementary Materials), and they did not depend on how long the strategies took to implement. The findings have theoretical implications for children's understanding of mathematical equivalence, as well as for their strategy use and metacognitive knowledge more broadly.

Implications for Research on Mathematical Equivalence

Consistent with previous research, elementary school children struggled to solve mathematical equivalence problems correctly (e.g., Fyfe & Rittle-Johnson, 2017; McNeil & Alibali, 2005). Success rates across the two studies were close to 50%, which is far from proficient. Other research has reported even lower performance on these problems, with lower bound success rates of 15–25% (e.g., Fyfe & Rittle-Johnson, 2017; McNeil et al., 2011). Direct comparisons with previous research are difficult given different inclusion

criteria; for example, Matthews and Rittle-Johnson (2009) worked with children from second through fifth grade and found a success rate on pretest equivalence problems of about 20%. However, that success rate was based on a subset of the original sample; they recruited 121 children and excluded 81 of them (67%) because they already scored higher than 50% on the pretest. These types of exclusions are common for research on interventions, in which the goal is to work with children who have room for improvement (e.g., Fyfe et al., 2012; Rittle-Johnson, 2006), but they make comparisons of baseline performance difficult across studies. In general, our results are consistent with the notion that students struggle to solve equivalence problems correctly, even when simpler items with operations on the right side only ($a = b + c$) are included.

Importantly, across all items, even the simpler ones, children used a variety of incorrect strategies (Matthews & Rittle-Johnson, 2009; McNeil & Alibali, 2005; Rittle-Johnson, 2006). Each incorrect strategy represented a faulty way of solving the problem. Yet, these incorrect strategies were not cognitively equivalent. Children were significantly more certain that they were correct when they used arithmetic-specific strategies than when they used other incorrect strategies. Why does this discrepancy in children's certainty ratings matter?

First, this discrepancy is uniquely predicted by the change-resistance account, and it provides support for this theory. According to the change-resistance theory (McNeil, 2014), children's difficulties with mathematical equivalence are not about something they lack (e.g., ability, working memory capacity); rather, it is about the *presence* of misconceptions that stem from the overgeneralization of strategies that are familiar and work well in arithmetic. One tacit assumption is that children are confident that the arithmetic-based strategies will work; they do not represent uncertain in-the-moment guesses. Although prior work has documented the *prevalence* of add-all and add-to-equal strategies (e.g., Perry et al., 1988), the current research is the first to demonstrate that children's *confidence* is uniquely high on these arithmetic-specific strategies relative to other incorrect strategies. Outside of the change-resistance account, there are no reasons to expect specific differences in certainty ratings as a function of the incorrect strategy type. For example, it is not the case that these arithmetic-specific strategies were faster to implement than other incorrect strategies, which could also lead to higher certainty. Thus, the findings lend support to the change-resistance account, which helps explain how early learning experiences (e.g., with traditional arithmetic) can influence cognitive development.

An open question concerns the causal relation between entrenchment and certainty. Here, certainty ratings were used as an indicator of entrenchment (see McNeil & Alibali, 2005), but bidirectional relations between confidence and entrenchment may be possible. Entrenched strategies that are used over and over again may produce greater confidence (e.g., due to familiarity, ease of retrieval and implementation). However, confidence may also causally contribute to entrenchment. Children may build up confidence in the *add-all* strategy when it works on a variety of traditional arithmetic problems (e.g., with two addends and three addends, with small numbers versus big numbers), and this confidence may cause them to use that strategy on novel problems and be resistant to using new strategies. Longitudinal relations between confidence and entrenchment should be examined in future research.

Second, this discrepancy in certainty ratings sheds light on the *persistence* of children's difficulties with mathematical equivalence. Misconceptions about mathematical equivalence are not easy to overcome. For example, although some children might learn a correct problem-solving strategy from a brief training, their knowledge change is often temporary (e.g., Cook et al., 2008). Further, misconceptions about equivalence can persist well beyond elementary school and into adulthood (e.g., Fyfe et al., 2020; McNeil et al., 2010). Children's overconfidence in their arithmetic-specific strategies may serve as a barrier to meaningful knowledge change. For example, Brown and Alibali (2018) found that feedback on children's equivalence performance was less likely to promote strategy change for children with high certainty in their initial strategies relative to children with low certainty in their initial strategies.

Third, this discrepancy in certainty ratings suggests the *specific* ways in which answers are wrong can matter (e.g., Congdon et al., 2018; Fitzsimmons et al., 2020; Fyfe et al., 2020; McNeil & Alibali, 2005). For example, Byrd and colleagues (2015) measured third-graders' and fifth-graders' understanding of the equal sign at the beginning of the school year. Some children provided correct, relational definitions of the equal sign (e.g., it means the same amount), and this was positively related to their end-of-year algebra performance. However, among the children who provided non-relational interpretations of the equal sign, some of them provided arithmetic-specific definitions (e.g., it means when you add something) and some provided other incorrect definitions (e.g., it means the end of the question). Children who provided arithmetic-specific definitions scored lower on the end-of-year measure compared to children who provided other incorrect definitions. These results are consistent with the change-resistance account; children's difficulties are not strictly about the lack of correct knowledge, but the presence of specific misconceptions (McNeil & Alibali, 2005).

Implications beyond Mathematical Equivalence

Although the current study focuses on mathematical equivalence, notions of entrenchment and change-resistance can be applied to a wide variety of domains in which learners hold misconceptions that are hard to change. In fact, the change-resistance account of mathematical equivalence stems from broader top-down approaches to cognition that help explain how early learning experiences can constrain later learning (see McNeil & Alibali, 2005). It also stems, in part, from classic research on mental sets, the common tendency to rely on familiar, well-practiced approaches, as demonstrated in Luchins' (1942) classic water jar experiments. These experiments showed participants were resistant to use new, more efficient strategies after getting used to using a different, less efficient strategy. The current findings may have implications for other domains in which misconceptions are resistant to change. For example, a large literature focuses on the whole number bias in fraction learning. The whole number bias refers to the fact that learners misapply their knowledge of whole numbers to non-whole numbers (i.e., decimals or fractions), such as the belief that numbers with more digits are larger (e.g., thinking that 1.45 is larger than 1.5; Roell et al., 2017). Findings from the current study suggest that children's confidence in strategies resulting from misconceptions may provide important insights into why these strategies are resistant to change. Additionally, they suggest that a promising avenue for

future research would be to determine if children's confidence in incorrect strategies stemming from misconceptions is higher than for other types of incorrect strategies in other domains with persistent misconceptions.

The current study also contributes to the literature on children's metacognitive abilities more broadly. First, the results of the current study corroborate previous findings that children tend to overestimate their knowledge during problem solving (e.g., Rinne & Mazzocco, 2014; Roebers & Spiess, 2017). In Study 1 and Study 2, children's average certainty ratings were very high, even on incorrect trials. For example, in Study 2, the most common incorrect strategy was *add-all* and children's average certainty ratings were 3.07 out of 4.00, indicating that most children were certain it was correct. Thus, the current findings suggest children tend to be overconfident in their own knowledge.

Second, the present results suggest that research in metacognition needs to move beyond correct versus incorrect knowledge to consider nuances in the ways in which knowledge can be wrong. In both studies, we found evidence that children were more confident when using arithmetic strategies than other incorrect strategies, despite the fact that both types of strategies represented faulty knowledge. Examining these nuances in strategy type can provide further insight into the reasons for faulty knowledge and into the different contexts in which metacognitive monitoring may be weak or strong. More broadly, it also suggests that certainty ratings may be a useful way to learn more about children's strategy use, especially in domains where students hold misconceptions (e.g., Durkin & Rittle-Johnson, 2015).

Future research should explore whether it is possible to effectively target children's certainty on specific strategies during training. Current approaches to facilitating knowledge of mathematical equivalence typically focus on modifying traditional arithmetic practice (e.g., McNeil et al., 2015) or providing instruction on correct concepts and strategies (e.g., Matthews & Rittle-Johnson, 2009; Perry, 1991). However, other research suggests a focus on incorrect strategies can be fruitful as well (e.g., Durkin & Rittle-Johnson, 2012). For example, middle school algebra students benefitted from explaining worked examples of quadratic equations, and the benefits were most robust when the examples displayed common incorrect strategies (Barbieri & Booth, 2020). Children learning about mathematical equivalence might benefit from studying clearly marked incorrect examples that showcase the arithmetic-specific strategies and from explaining why students might believe they are correct. This approach would target the specific incorrect strategies that children tend to use, and potentially help children understand why their certainty may be misplaced.

Limitations and Future Directions

Limitations of the current study suggest additional directions for future research. For example, the analyses were based on data from small samples of children in a fairly narrow age range. Future research should include larger samples with children spanning multiple ages to better track the development of children's mathematical equivalence knowledge and their certainty on various strategies. Also, the generalizability of the conclusions should be tested in samples with varying educational experiences and demographic characteristics.

Although misconceptions about mathematical equivalence have been documented in children from a variety of countries (e.g., Humberstone & Reeve, 2008; Molina et al., 2009), these misconceptions are not universal (e.g., Li et al., 2008).

Additional limitations concern the measures used in the current study. Children solved a small set of mathematical equivalence problems. Including a larger, more varied set of items would allow for more fine-grained analyses across different features of the problems and across the strategy types. Additionally, we characterized children's strategy use based on their numerical responses, which limited our analysis to different types of incorrect strategies. If we had gathered verbal reports, we would have been able to examine different types of correct strategies as well. Further, response time data was only available in Study 1 with a small sample of children. The conclusions are also dependent on the way in which certainty was measured. In a recent review, Desoete and De Craene (2019) highlighted the large number of diverse measures used to assess different aspects of metacognition and that "there still is no 'gold standard'" (p. 571). The certainty rating measure in the current research was selected based on prior work within this age range (e.g., Desoete et al., 2006), but future research is needed to test the generalizability of our conclusions when different measures are used.

Conclusion

Despite these limitations, we provided novel evidence in support of the change-resistance account by considering children's metacognitive abilities. Children struggled to solve mathematical equivalence problems correctly and frequently used strategies that stem from their narrow experience with arithmetic. Critically, children were more certain that they were correct when they employed these arithmetic-specific strategies relative to other incorrect strategies, indicating that not all incorrect strategies were cognitively equivalent. These results shed light on the development of misconceptions in early childhood as well as on children's metacognitive abilities. Going beyond accuracy – and focusing on the use of specific strategies – is a critical mechanism for understanding cognitive development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Nelson and Grenell were supported by a training grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Number T32HD007475. The content is the responsibility of the authors and does not represent the official views of the National Institutes of Health. The authors thank the principals, teachers, and students at participating schools.

References

- Alibali MW (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, 35, 127–145. 10.1037/0012-1649.35.1.127. [PubMed: 9923470]
- Alibali MW, Brown SA, & Menendez D (2019). Understanding strategy change: Contextual, individual, and metacognitive factors. *Advances in Child Development and Behavior*, 56, 227–256. 10.1016/bs.acdb.2018.11.004 [PubMed: 30846048]

- Barbieri CA, & Booth JL (2020). Mistakes on display: Incorrect examples refine equation solving and algebraic feature knowledge. *Applied Cognitive Psychology*, 34 (4), 862–878. 10.1002/acp.3663
- Baroody AJ, & Ginsburg HP (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal*, 84, 199–212. 10.1086/461356
- Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. 10.18637/jss.v067.i01
- Behr M, Erlwanger S, & Nichols E (1980). How children view the equal sign. *Mathematics Teaching*, 92, 13–15.
- Bellon E, Fias W, & De Smedt B (2019). More than number sense: The additional role of executive functions and metacognition in arithmetic. *Journal of Experimental Child Psychology*, 182, 38–60. 10.1016/j.jecp.2019.01.012 [PubMed: 30807905]
- Bjorklund DF, & Green BL (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, 47(1), 46–54.
- Brown SA, & Alibali MW (2018). Promoting strategy change: Mere exposure to alternative strategies helps, but feedback can hurt. *Journal of Cognition and Development*, 19(3), 301–324. 10.1080/15248372.2018.1477778
- Byrd CE, McNeil NM, Chesney DL, & Matthews PG (2015). A specific misconception of the equal sign acts as a barrier to children's learning of early algebra. *Learning and Individual Differences*, 38, 61–67. 10.1016/j.lindif.2015.01.001
- Fyfe ER, Byers C, & Nelson LJ (2021). The benefits of a metacognitive lesson on children's understanding of mathematical equivalence, arithmetic, and place value. *Journal of Educational Psychology*. Advance online publication. 10.1037/edu0000715.
- Carpenter TP, Franke ML, & Levi L (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Heinemann.
- Congdon EL, Kwon MK, & Levine SC (2018). Learning to measure through action and gesture: Children's prior knowledge matters. *Cognition*, 180, 182–190. 10.1016/j.cognition.2018.07.002 [PubMed: 30055339]
- Cook SW, Mitchell Z, & Goldin-Meadow S (2008). Gesturing makes learning last. *Cognition*, 106, 1047–1058. 10.1016/j.cognition.2007.04.010 [PubMed: 17560971]
- Coughlin C, Hembacher E, Lyons KE, & Ghetti S (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, 18, 957–971. 10.1111/desc.12271 [PubMed: 25483236]
- De Clercq A, Desoete A, & Roeyers H (2000). EPA2000: a multilingual, programmable computer assessment of off-line metacognition in children with mathematical-learning disabilities. *Behavior Research Methods, Instruments, & Computers*, 32, 304–311. 10.3758/BF03207799
- Desoete A, & De Craene B (2019). Metacognition and mathematics education: An overview. *ZDM*, 51, 565–575. 10.1007/s11858-019-01060-w
- Desoete A, Roeyers H, & Huylebroeck A (2006). Metacognitive skills in Belgian third grade children (age 8 to 9) with and without mathematical learning disabilities. *Metacognition and Learning*, 1(2), 119–135. 10.1007/s11409-006-8152-9
- Destan N, & Roebbers CM (2015). What are the metacognitive costs of young children's overconfidence? *Metacognition Learning*, 10, 347–374. 10.1007/s11409-014-9133-z
- Durkin K, & Rittle-Johnson B (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, 22(3), 206–214. 10.1016/j.learninstruc.2011.11.001
- Durkin K, & Rittle-Johnson B (2015). Diagnosing misconceptions: Revealing changing decimal fraction knowledge. *Learning and Instruction*, 37, 21–29. 10.1016/j.learninstruc.2014.08.003
- Fazio LK, DeWolf M, & Siegler RS (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 1–16. 10.1037/xlm0000153
- Finn B, & Metcalfe J (2014). Overconfidence in children's multi-trial judgements of learning. *Learning and Instruction*, 32, 1–9. 10.1016/j.learninstruc.2014.01.001

- Fitzsimmons CJ, Thompson CA, & Sidney PG (2020). Do adults treat equivalent fractions equally? Adults' strategies and errors during fraction reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(11), 2049–2074. 10.1037/xlm0000839
- Flavell JH (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911. 10.1037/0003-066X.34.10.906
- Fyfe ER, Matthews PG, & Amsel E (2020). College developmental math students' knowledge of the equal sign. *Educational Studies in Mathematics*, 104, 65–85. 10.1007/s10649-020-09947-2
- Fyfe ER, & Rittle-Johnson B (2017). Mathematics practice without feedback: A desirable difficulty in a classroom setting. *Instructional Science*, 45, 177–194. 10.1007/s11251-016-9401-1
- Fyfe ER, Rittle-Johnson B, & DeCaro MS (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104, 1094–1108. 10.1037/a0028389
- Geary DC, Hoard MK, Byrd-Craven J, Nugent L, & Numtee C (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, 78, 1343–1359. 10.1111/j.1467-8624.2007.01069.x [PubMed: 17650142]
- Hornburg CB, Wang L, & McNeil NM (2018). Comparing meta-analysis and individual person data analysis using raw data on children's understanding of equivalence. *Child Development*, 89, 1983–1995. 10.1111/cdev.13058 [PubMed: 29574694]
- Humberstone J, & Reeve RA (2008). Profiles of algebraic competence. *Learning and Instruction*, 18, 354–367. 10.1016/j.learninstruc.2007.07.002
- Kieran C (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, 12, 317–326. 10.1007/BF00311062
- Knuth EJ, Stephens AC, McNeil NM, & Alibali MW (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, 37, 297–312.
- Koriat A, Ma'ayan H, & Nussinson R (2006). The intricate relationships between monitoring and control in metacognition: lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, 135, 36–69. 10.1037/0096-3445.135.1.36 [PubMed: 16478315]
- Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. 10.18637/jss.v082.i13
- Lemaire P (2017). *Cognitive Development from a Strategy Perspective*. NY: Routledge.
- Li X, Ding M, Capraro MM, & Capraro RM (2008). Sources of differences in children's understanding of mathematical equality. *Cognition and Instruction*, 26, 1–23. 10.1080/07370000801980845
- Luchins AS (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, 54(6), i–, 95. 10.1037/h0093502
- Matthews PG, & Fuchs LS (2020). Keys to the gate? Equal sign knowledge at second grade predicts fourth grade algebra competence. *Child Development*, 91, 14–28. 10.1111/cdev.13144
- Matthews P, & Rittle-Johnson B (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, 104, 1–21. 10.1016/j.jecp.2008.08.004 [PubMed: 18937952]
- McGilly K, & Siegler RS (1990). The influence of encoding and strategic knowledge on children's choices among serial recall strategies. *Developmental Psychology*, 26(6), 931–941. 10.1037/0012-1649.26.6.931
- McNeil NM (2014). A change–resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, 8, 42–47. 10.1111/cdep.12062
- McNeil NM (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, 43(3), 687–695. 10.1037/0012-1649.43.3.687 [PubMed: 17484580]
- McNeil NM, & Alibali MW (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, 76, 883–899. 10.1111/j.1467-8624.2005.00884.x [PubMed: 16026503]
- McNeil NM, Fyfe ER, Petersen LA, Dunwiddie AE, & Brletic-Shipley H (2011). Benefits of practicing $4 = 2 + 2$: Nontraditional problem formats facilitate children's

- understanding of mathematical equivalence. *Child Development*, 82(5), 1620–1633. 10.1111/j.1467-8624.2011.01622.x [PubMed: 21793817]
- McNeil NM, Fyfe ER, & Dunwiddie AE (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology*, 107(2), 423–436. 10.1037/a0037687
- McNeil NM, Grandau L, Knuth EJ, Alibali MW, Stephens AC, Hattikudur S, & Krill DE (2006). Middle-school students' understanding of the equal sign: The books they read can't help. *Cognition and Instruction*, 24, 367–385. 10.1207/s1532690xci2403_3
- McNeil NM, Hornburg CB, Devlin BL, Carrazza C, & McKeever MO (2019). Consequences of individual differences in children's formal understanding of mathematical equivalence. *Child Development*, 90(3), 940–956. 10.1111/cdev.12948 [PubMed: 28902386]
- McNeil NM, Rittle-Johnson B, Hattikudur S, & Petersen LA (2010). Continuity in representation between children and adults: Arithmetic knowledge hinders undergraduates' algebraic problem solving. *Journal of Cognition and Development*, 11, 437–457. 10.1080/15248372.2010.516421
- Molina M, Castro E, & Castro E (2009). Elementary students' understanding of the equal sign in number sentences. *Education & Psychology*, 17, 341–368.
- Nelson LJ, & Fyfe ER (2019N). Metacognitive monitoring and help-seeking decisions on mathematical equivalence problems. *Metacognition and Learning*, 14, 167–187. 10.1007/s11409-019-09203-w.
- Nelson TO, & Narens L (1990). Metamemory: A theoretical framework and new findings. In Bower INGH (Ed.), *The psychology of learning and instruction: Advances in Research and Theory* (vol. 26, pp. 125–141). New York, NY: Academic Press.
- Perry M (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development*, 6, 449–468. 10.1016/0885-2014(91)90049-J
- Perry M, Church RB, & Goldin-Meadow S (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, 3, 359–400. 10.1016/0885-2014(88)90021-4
- Powell SR (2012). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal*, 112, 627–648. 10.1086/665009 [PubMed: 22707760]
- R Core Team, 2020, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (<https://www.R-project.org/>).
- Rinne LF, & Mazzocco MM (2014). Knowing right from wrong in mental arithmetic judgments: Calibration of confidence predicts the development of accuracy. *PloS one*, 9(7), 98663. 10.1371/journal.pone.0098663
- Rittle-Johnson B (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, 77, 1–15. 10.1111/j.1467-8624.2006.00852.x [PubMed: 16460521]
- Rittle-Johnson B, Matthews PG, Taylor RS, & McEldoon KL (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, 103(1), 85–104. 10.1037/a0021334
- Rittle-Johnson B, Star JR, & Durkin K (2020). How can cognitive-science research help improve education? The case of comparing multiple strategies to improve mathematics learning and teaching. *Current Directions in Psychological Science*, 29, 599–609. 10.1177/0963721420969365
- Roebers CM, & Spiess M (2017). The development of metacognitive monitoring and control in second graders: A short-term longitudinal study. *Journal of Cognition and Development*, 18(1), 110–128. 10.1080/15248372.2016.1157079
- Roell M, Viarouge A, Houde O, & Borst G (2017). Inhibitory control and decimal number comparison in school-aged children. *PLOS ONE*, 1–17. 10.1371/journal.pone.0188276
- Siegler RS (2000). The rebirth of children's learning. *Child Development*, 71, 26–35. 10.1111/1467-8624.00115 [PubMed: 10836555]
- Siegler RS (1983). Five generalizations about cognitive development. *American Psychologist*, 38(3), 263–277.
- Stephens A, Knuth E, Blanton M, Isler I, Gardiner A, & Marum T (2013). Equation structure and the meaning of the equal sign: The impact of task selection in eliciting elementary students' understandings. *The Journal of Mathematical Behavior*, 32(2), 173–182. 10.1016/j.jmathb.2013.02.001

Vo VA, Li R, Kornell N, Pouget A, & Cantlon JF (2014). Young children bet on their numerical skills: metacognition in the numerical domain. *Psychological Science*, 25(9), 1712–1721. 10.1177/0956797614538 [PubMed: 24973137]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Children varied in their strategy use and certainty on math equivalence problems.
- Arithmetic-specific strategies were the most common type of incorrect strategy.
- Children's metacognitive abilities were modest due to overconfidence.
- Children's confidence varied depending on the type of incorrect strategy used.
- Results are consistent with the change-resistance theory.

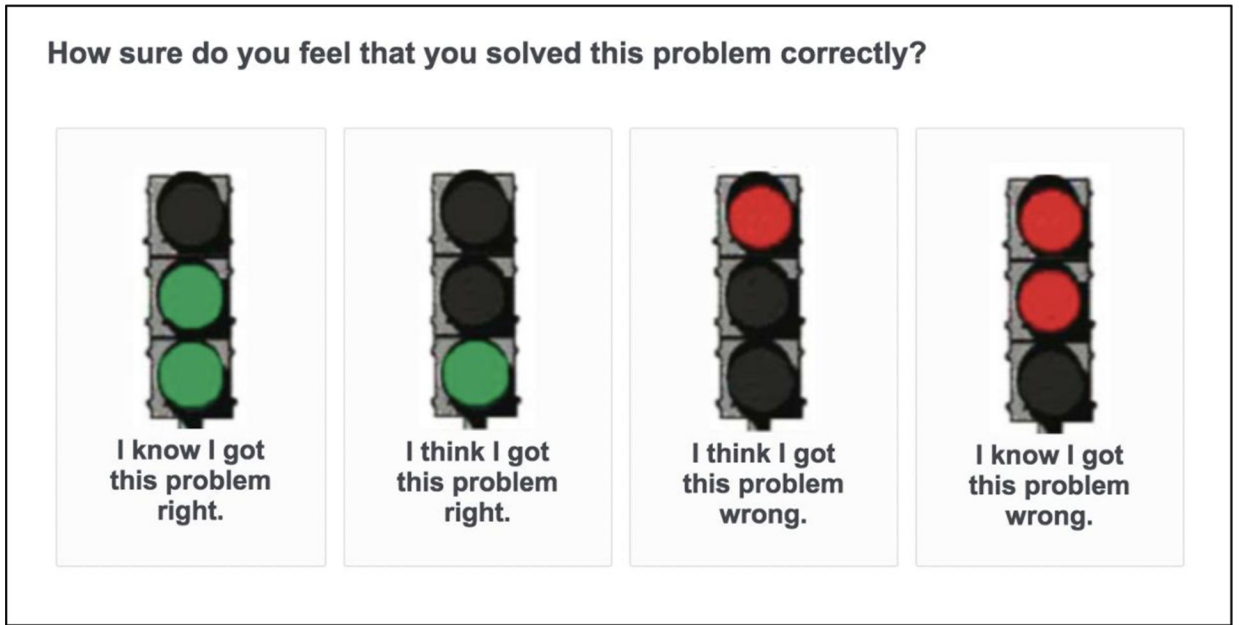


Figure 1.
Certainty Rating Scale

Table 1

Mathematical Equivalence Problems Presented to Children in Study 1 and Study 2

Study 1			Study 2	
Item		Mean Percent Correct	Item	Mean Percent Correct
<i>Set 1</i>	<i>Set 2</i>			
$7 = _ + 4$	$8 = 3 + _$	58%	$8 = _ + 3$	92%
$3 + 7 = 3 + _$	$4 + 6 = 4 + _$	48%	$3 + 7 + 6 = _ + 6$	38%
$9 + 4 = 3 + _$	$5 + 9 = 4 + _$	40%	$2 + 7 = 6 + _$	46%
$3 + 7 + 8 = _ + 8$	$8 + 3 + 7 = _ + 7$	44%	$7 = 4 + _$	80%
$5 + 3 + 4 = 5 + _$	$4 + 5 + 3 = 4 + _$	46%	$3 + 4 = _ + 5$	49%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Children's Strategies on Mathematical Equivalence Problems

Strategy Type	Description	Example Responses
Correct	Made both sides of the equal sign the same amount	$3 + 7 + 8 = 10 + 8$
Arithmetic-Incorrect		
Add-All	Added all the numbers in the problem	$3 + 7 + 8 = 26 + 8$
Add-to-Equal	Added all the numbers before the equal sign	$3 + 7 + 8 = 18 + 8$
Other-Incorrect		
Carry	Copied one number from the problem	$3 + 7 + 8 = 3 + 8$
Idiosyncratic	Idiosyncratic combination of numbers	$3 + 7 + 8 = 16 + 8$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Frequency of Certainty Ratings Across Problems

	Certainty Rating	Frequency	Percentage
Study 1	1	5	2%
	2	37	14%
	3	111	43%
	4	107	41%
Study 2	1	86	10%
	2	103	12%
	3	216	24%
	4	473	54%
	Missing	4	1%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Frequency of Strategy use and Average Certainty Rating by Strategy Type Across Problems

	Strategy Type	Strategy	Frequency	Percentage	Certainty Rating <i>M</i> (<i>SD</i>)	Response Time <i>M</i> (<i>SD</i>)
Study 1	Correct	Correct	123	47%	3.42 (.64)	32.12 (18.52)
	Arithmetic	Add-All	63	24%	3.17 (.81)	40.83 (18.35)
		Add-to-Equal	39	15%	3.10 (.72)	60.82 (39.18)
	Other-Incorrect	Carry	14	5%	3.00 (.88)	52.97 (30.02)
		Idiosyncratic	20	8%	2.75 (.85)	61.50 (34.72)
	No Answer	Blank	1	.4%	1.00 (—)	95.46 (—)
Study 2	Correct	Correct	511	58%	3.63 (.71)	—
	Arithmetic	Add-All	132	15%	3.07 (.96)	—
		Add-to-Equal	83	9%	2.79 (.94)	—
	Other-Incorrect	Carry	16	2%	2.63 (.89)	—
		Idiosyncratic	99	11%	2.55 (1.00)	—
	No Answer	Blank	41	5%	1.39 (.86)	—

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Results for Mixed-Effects Regression Models Predicting Children's Certainty Ratings for Study 1

Variable	<i>B</i> (<i>SE</i>)	<i>t</i>	95% <i>CI</i>	<i>p</i>
Model 1a				
Contrast 1: Arithmetic vs. Correct	0.03 (0.11)	0.26	[-.22, .26]	0.79
Contrast 2: Arithmetic vs. Other-Incorrect	-0.53 (0.14)	-3.81	[-.82, -.26]	<.001
Model 1b				
Response Time	-0.004 (0.001)	-2.87	[-.01, -.001]	0.005
Contrast 1: Arithmetic vs. Correct	-0.02 (0.11)	-0.13	[-.27, .21]	0.89
Contrast 2: Arithmetic vs. Other-Incorrect	-0.49 (0.14)	-3.55	[-.78, -.21]	<.001
Model 2a				
Contrast 1: Arithmetic vs. Correct	-0.02 (0.14)	-0.14	[-.31, .26]	0.89
Contrast 2: Arithmetic vs. Other-Incorrect	-0.54 (0.15)	-3.58	[-.85, -.24]	<.001
Model 2b				
Response Time	-0.002 (0.002)	-1.50	[-.01, .001]	0.14
Contrast 1: Arithmetic vs. Correct	-0.05 (0.14)	-0.37	[-.34, .23]	0.71
Contrast 2: Arithmetic vs. Other-Incorrect	-0.51 (0.15)	-3.38	[-.82, -.21]	<.001

Note. Model 1a and Model 1b included all items. Model 2a and Model 2b included a subset of items as the easier items (i.e., problems with operations on the right side only) were removed. Strategy Type was dummy-coded with the "Arithmetic Strategies" as the reference group.

Table 6

Results for Mixed-Effects Regression Models for Study 2

Variable	<i>B</i> (<i>SE</i>)	<i>t</i>	95% <i>CI</i>	<i>p</i>
Model 1				
Contrast 1: Arithmetic vs. Correct	44 (.07)	6.27	[.30, .58]	<.001
Contrast 2: Arithmetic vs. Other-Incorrect	-.35(.09)	-3.94	[-.53, -.018]	<.001
Model 2				
Contrast 1: Arithmetic vs. Correct	.34(.09)	3.76	[.15, .52]	<.001
Contrast 2: Arithmetic vs. Other-Incorrect	-.39(.10)	-3.87	[-.59, -.19]	<.001

Note. Model 1 included all items and Model 2 included a subset of items as the easier items (i.e., problems with operations on the right side only) were removed. Strategy Type was dummy-coded with the “Arithmetic Strategies” as the reference group.