It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence

Nicole M. McNeil, Dana L. Chesney, Percival G. Matthews, Emily R. Fyfe, Lori A. Petersen, April E. Dunwiddie, and Mary C. Wheeler

University of Notre Dame

Final version submitted to *Journal of Educational Psychology*, May 14, 2012

Author Note

Nicole M. McNeil, Dana L. Chesney, Percival G. Matthews, Emily R. Fyfe, Lori A. Petersen, April E. Dunwiddie, and Mary C. Wheeler, Department of Psychology, University of Notre Dame. Emily R. Fyfe is now at the Department of Psychology and Human Development, Vanderbilt University.

This research was supported by Grant R305B070297 from the U.S. Department of Education, Institute of Education Sciences. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Thanks to Martha Alibali and Heather Brletic-Shipley for their input. Thanks also to Noelle Crooks, Matt Gibson, Jenny Heil, Zachary Montez, Elizabeth Nye, Erica Pepitone, Joanna Thurnes, Lauren Wester, and Krysten Williams for help with coding and tutoring. This research would not have been possible without the support of many administrators, teachers, parents, and students in the greater South Bend area.

Correspondence concerning this article should be addressed to Nicole M. McNeil, Department of Psychology, University of Notre Dame, 118 Haggar Hall, Notre Dame, IN, 46556; email: nmcneil@nd.edu.

Abstract

This experiment tested the hypothesis that organizing arithmetic fact practice by equivalent values facilitates children's understanding of math equivalence. Children (*M* age = 8;6, *N* = 104) were randomly assigned to one of three practice conditions: (a) equivalent values, in which problems were grouped by equivalent sums (e.g., 3 + 4 = 7, 2 + 5 = 7, etc.), (b) iterative, in which problems were grouped iteratively by shared addend (e.g., 3 + 1 = 4, 3 + 2 = 5, etc.), or (c) no extra practice, in which children did not receive any practice over and above what they ordinarily receive at school and home. Children then completed measures to assess their understanding of math equivalence. Children who practiced facts organized by equivalent values demonstrated a better understanding of math equivalence than children in the other two conditions. Results suggest that organizing arithmetic facts into conceptually related groupings may help children improve their understanding of math equivalence.

It pays to be organized: Organizing arithmetic knowledge around equivalent values facilitates understanding of math equivalence

Researchers often study macro-level differences in children's environments and how they affect children's learning and development (e.g., Bradley & Corwyn, 2002; Jordan, Huttenlocher, & Levine, 1992; Tran & Weinraub, 2006). However, a growing body of research suggests that even relatively minor differences in the structure of children's early input can play a central role in shaping and constraining children's understanding of fundamental concepts (e.g., Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; McNeil, Fyfe, Petersen, Dunwiddie, Brletic-Shipley, 2011; Ramani & Siegler, 2008). These studies suggest that parents and teachers must consider the specific input they provide to children and how the structure of that input maps onto the underlying structure of the conceptual domain they want children to learn. Indeed, as Gelman and Williams (1998) have cogently argued, the structure of one's learning environment can determine the structure of one's conceptual knowledge and its ultimate compatibility with the target domain. In the present study, we examined how a seemingly minor modification to the structure of arithmetic practice affects children's understanding of mathematical equivalence.

Mathematical equivalence is one of the most fundamental concepts in mathematics. Formally represented by the equal sign, it is the principle that two sides of an equation are interchangeable or represent the same value. Researchers and educators alike agree that math equivalence is one of the most important concepts for developing children's early algebraic thinking (Falkner, Levi, & Carpenter, 1999; Jacobs, Franke, Carpenter, Levi, & Battey, 2007; Kieran, 1992; Knuth, Stephens, McNeil, & Alibali, 2006; National Research Council, 2001; Steinberg, Sleeman & Ktorza, 1990). Moreover, because algebra serves as a gateway not only into higher mathematics but also into higher education more generally, building an early understanding of math equivalence may have positive consequences for children's long-term academic success (Adelman, 2006; Blanton & Kaput, 2005; Moses & Cobb, 2001; National Research Council, 2001; NCTM, 2000).

Unfortunately, most children in the U.S. develop a poor understanding of math equivalence (e.g., Alibali, 1999; Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols, 1980; Falkner et al., 1999; Jacobs et al., 2007; Li, Ding, Capraro, & Capraro, 2008; McNeil, 2008; Perry, 1991; Powell & Fuchs, 2010; Renwick, 1932; Weaver, 1973). This is most typically evidenced by children's failure to solve what are generally termed "mathematical equivalence problems" – open equations with operations on both sides of the equal sign (e.g., $3 + 7 + 5 = 3 + \_\_$). For example, across nine studies, McNeil (2005) found that the vast majority (about 80%) of children in the U.S. aged 7-11 failed to solve these problems correctly.

Historically, some researchers have attributed children's difficulties with math equivalence to domain general conceptual limitations in childhood (Collis, 1974, as cited in Kieran, 1980; Piaget & Szeminska, 1995/1941), or to developmental changes in some parameter of the working memory system such as total capacity, efficiency, or processing speed (e.g., Case, 1978). These researchers posit that children's misunderstandings of math equivalence result from some cognitive deficit that children have relative to adults. However, a growing body of research suggests that children's difficulties are largely due to children's early experiences with mathematics (e.g., Baroody & Ginsburg, 1983; Li et al., 2008; McNeil, 2008; McNeil & Alibali, 2005b; Seo & Ginsburg, 2003). This view is typified by McNeil and Alibali's (2005b) change-resistance account, which suggests that children's difficulties with math equivalence stem from the extraction and entrenchment of patterns children routinely encounter in their formal experiences with arithmetic.

The majority of children in the U.S. learn arithmetic in a very procedural fashion. Moreover, arithmetic problems are almost always presented in a traditional format with the operations to the left of the equal sign and the "answer" to the right (e.g., 3 + 4 = 7, McNeil et al., 2006; Seo & Ginsburg, 2003). Overuse of this format to the exclusion of others fails to highlight the interchangeable nature of the two sides of an equation (McNeil et al., 2011; Weaver, 1973). As a result of this overly narrow experience, children often extract three *operational patterns* that do not generalize beyond arithmetic (McNeil & Alibali, 2005b). First, children learn a perceptual pattern: the "operations on left side" format (Alibali, Phillips, & Fischer, 2009; Cobb, 1987; McNeil & Alibali, 2004). Second, children learn a default problem-solving strategy, which is to "perform all given operations on all given numbers" (McNeil & Alibali, 2005b; Perry, Church, & Goldin-Meadow, 1988). Finally, children interpret the equal sign as a "do something" symbol that means "add up all the numbers" (Baroody & Ginsburg, 1983; Behr et al., 1980; Kieran, 1981).

Reliance on these operational patterns often persists for many years, serving as an obstacle to the construction of flexible problem solving skills and algebraic reasoning (Knuth et al., 2006; McNeil et al., 2006; McNeil, Rittle-Johnson, Hattikudur, & Petersen, 2010; Steinberg et al., 1990). As a result, most elementary school children (ages 7-11) solve problems such as 7 + 4 + 5 = 7 + __ incorrectly by adding all the numbers and writing 23 in the blank (McNeil & Alibali, 2005b). Furthermore, when asked to reconstruct problems such as "7 + 4 + 5 = 7 + __" after viewing them briefly, many children rely on the "operations on left side" pattern and write "7 + 4 + 5 + 7 = __" (Alibali et al., 2009; McNeil & Alibali, 2004). Indeed, most children actually reject closed equations that do not take a standard $a + b = c$ format, considering equations such as 3 = 3 and 7 + 6 = 6 + 6 + 1 to be false or nonsensical (Baroody & Ginsburg, 1983; Behr et. al., 1980; Falkner et al, 1999). Finally, few children interpret the equal sign as a *relational* symbol that asserts a proposition about the interchangeability of the two sides of an equation (Baroody & Ginsburg, 1983; McNeil & Alibali, 2005a).

Although most children in the U.S. exhibit these misconceptions, a substantial minority constructs a good understanding of math equivalence, despite receiving the same formal experiences with arithmetic. For example, in each of the aforementioned studies approximately 10-25% of children solved math equivalence problems correctly, reconstructed the problems correctly after viewing them briefly, and defined the equal sign relationally. These "successful" children attended the same schools, learned from the same teachers and textbooks, came from families with comparable socio-economic status, and even scored similarly on tests of computational fluency as children who were "unsuccessful." What factors led these children to extract the appropriate patterns from their formal experiences with arithmetic? Why did they not acquire entrenched representations of the operational patterns like most other children did?

At first glance, one may simply assume that the successful children are just "smarter" or "better at mathematics overall." However, that position is untenable, given that researchers have not found strong positive correlations between understanding of math equivalence and performance on tests of computational fluency. They also have not found associations between understanding of math equivalence and other proxies of general competence, such as age or performance on standardized achievement tests. In fact, within some age ranges (e.g., 7-9) younger children actually perform better solving math equivalence problems than do older children (McNeil, 2007). Thus, many of the general factors that are typically associated with the development of math cognition do not appear to play a leading role in children's understanding of math equivalence. Instead, specific aspects of domain knowledge may play a greater role.

One aspect of domain knowledge that may affect children's understanding of math equivalence is the way in which children's knowledge of arithmetic facts is organized. Decades of research have demonstrated that the organization of a person's knowledge is at least as influential on learning and performance in a given domain as the amount of knowledge that person might possess (e.g. Chase & Simon, 1973; Lindberg, 1980; see Chi & Ceci, 1987 for a review). Thus, we should expect the organization of children's knowledge of arithmetic facts to have profound effects on what they learn from their formal experiences with arithmetic. Consistent with this view, Gelman & Williams (1998) suggested that children who organize their knowledge of arithmetic such that they understand equivalent ways to state a math problem (e.g., 2 + 5 = 7 and 1 + 1 + 1 + 1 + 1 + 1 + 1 = 7) have a redundancy in their framework of numerical concepts that allows them to map their learning of arithmetic facts onto arithmetic principles. That is, having several different ways to represent the number "7" may help children construct a deeper understanding both of the number "7" and of arithmetic principles such as the properties of addition (e.g., commutative, associative, additive identity) that allow the production of multiple equivalent representations. An understanding of the properties of addition provides the basis for constructing an algebraic understanding of math equivalence (NCISLA, 2000).

Following this logic, we predicted that arithmetic practice that is designed to help children mentally organize arithmetic facts in conceptually related groupings leads children to gain a better understanding of math equivalence than practice that is not designed for this purpose. We were specifically interested in practice designed to help children organize their knowledge of arithmetic around number combinations that yield equivalent sums (e.g., 1 + 6, 4 +

3, 1 + 3 + 3, etc.). There are at least three ways in which this type of organized practice may facilitate understanding of math equivalence. First, as mentioned above, it may highlight that there are many different ways to represent a number and help children construct an understanding of the addition principles that are the basis for understanding math equivalence. Second, when problems with equivalent sums are practiced in succession, addend pairs that are equal to one another (e.g., 3 + 4 and 5 + 2) become increasingly inter-associated. Subsequently, when faced with an equation that involves one or more of those addend pairs (e.g., 3 + 4 = __ + 2), children who have practiced with this organization may see 3 + 4 and immediately activate other equivalent addend pairs, including 5 + 2. The concurrent activation of both 3 + 4 and 5 + 2 may then operate via top-down processes to influence encoding of, strategy selection on, and/or conceptualization of equations involving these facts (cf. Bruner, 1957; McGilly & Siegler, 1990; Rumelhart, 1980). Third, practice with problems organized by equivalent sums may enable children to induce equivalent added pairs based on transitive relations (e.g., if 3 + 4 = 7 and 5 + 2 = 7, then 3 + 4 = 5 + 2). This particular insight—that an addend pair can be equivalent not only to a single number, but also to other addend pairs—is a key milestone in the construction of an understanding of math equivalence (Behr et. al., 1980; Kieran, 1981). Thus, we hypothesize that arithmetic practice that is organized by equivalent sums will lead children to construct a better understanding of math equivalence than practice that is not organized by equivalent sums.

McNeil et al. (2010) found anecdotal support for this hypothesis. In the study, undergraduates who had learned arithmetic in the United States (in elementary school) exhibited a poorer understanding of math equivalence than did undergraduates who had learned arithmetic in countries in East Asia that consistently perform at the top in international comparisons of mathematics performance. The effect was large, and it was not due to general differences in mathematics aptitude as measured by scores on the quantitative section of college entrance exams (e.g., ACT and SAT). At the end of the study, McNeil et al. interviewed the undergraduates about their experiences with arithmetic in elementary school (see McNeil, 2009). A majority of the undergraduates who learned arithmetic in high-achieving countries remembered learning and practicing arithmetic organized by equivalent values (e.g., “all the addend pairs that equal 10”). In contrast, undergraduates who learned arithmetic in the U.S. remembered learning and practicing arithmetic problems organized in an iterative fashion like a traditional arithmetic table (i.e., all the ones, all the twos, all the threes, etc.). It is possible that this difference in the organization of arithmetic practice was one of the many factors contributing to the differences between groups in understanding of math equivalence.

The anecdote underscores our hypothesis that it is practice organized by equivalent values specifically—rather than organized practice in general—that should lead to improved understanding of math equivalence. Alternative organizational structures, such as the iterative practice identified by the undergraduates in McNeil et al.’s (2010) study, should not yield the same benefits as practice organized by equivalent sums. Although iterative practice does provide mental organization, it does not highlight equivalent values or the interchangeability of arithmetic expressions.

In the present study, we went beyond anecdotes and performed an experiment to test the effects of different organizations of arithmetic practice on children’s knowledge. We predicted that subtle changes to the ordering of addition facts practice could lead to improvements in understanding of math equivalence. Previous research has shown that well organized arithmetic facts practice can facilitate understanding that extends beyond simple retrieval of those facts. For example, Canobi (2009) found that organizing arithmetic practice in conceptual sequences, such as pairs of commuted addition problems, increased second graders' use of more sophisticated solution strategies, such as decomposition and retrieval (cf., Steinberg, 1985). Additionally, McNeil et al. (2011) found that children who received arithmetic practice with problems presented in a nontraditional format (e.g., 17 = 9 + 8) constructed a better understanding of math equivalence than children who received arithmetic practice with problems presented in the traditional format (e.g., 9 + 8 = 17). These results indicate that seemingly small changes to children's early arithmetic practice can help restructure knowledge in ways that support adaptive strategy use and conceptual understanding.

To test our hypothesis, we randomly assigned individual children to one of three conditions in which they received practice with addition facts organized by equivalent values, practice with addition facts organized in an iterative fashion, or no extra practice. Following the intervention (or no intervention for children in the no extra practice condition), we assessed children's understanding of math equivalence. After this assessment, we randomly assigned children in the no extra practice condition to one of the two practice conditions. Finally, all children received a follow-up assessment two weeks after their final practice session. Thus, we had two designs: the main experiment was a posttest-only randomized experiment with a follow-up, and the supplemental experiment performed with children in the no extra practice condition was a randomized experiment with a pretest, intervention, and follow-up. In both experiments, we expected children to construct a better understanding of math equivalence after practicing addition problems that were organized by equivalent values when compared to iterative or no extra practice conditions. Note that we expected the advantage conferred by organizing facts by equivalent sums to be great enough to improve understanding of math equivalence even though practice remained in the standard $a + b = c$ format, which has been implicated in the construction of misconceptions about math equivalence (McNeil & Alibali, 2005b).

Method

Participants

Participants were 108 children (ages 7-9) recruited from a diverse range of public and private elementary schools (approximately 62% received free or reduced price lunch, with schools ranging from 0-79% in terms of the percentage of children who qualified for free or reduced price lunch). Four children were excluded from the analysis because they did not complete one or more of the sessions. Thus, the final sample consisted of 104 children (*M* age = 8 yrs, 6 mo; 34 first semester second graders, 7 second semester second graders, 48 first semester third graders, 15 second semester third graders; 49 boys, 55 girls; 57% white, 26% African American or black, 12% Hispanic or Latino, 3% multiethnic, and 2% Asian) from a mid-sized city in the Midwestern United States. Although children with learning disabilities were not specifically excluded from participating in the study, none of the participating children were identified as having a learning disability.

Design

The design was a posttest-only randomized experiment, with random assignment at the individual level. Each child was randomly assigned to one of three conditions: (a) equivalent values (n = 36), in which practice problems were grouped and presented by equivalent values (e.g., 3 + 4 = __, 5 + 2 = __, 1 + 6 = __), (b) iterative (n = 34), in which practice problems were grouped and presented iteratively by shared addend, according to an addition table (e.g., 3 + 4 = __, 3 + 5 = __, 3 + 6 = __), or (c) no extra practice (n = 34), in which children did not receive any practice over and above what they ordinarily receive at school and home.

Although our use of random assignment ensured that any differences between conditions at the outset of the experiment could be attributed to chance, we nonetheless compared children in the three conditions to see if they differed in terms of their background characteristics. The conditions were generally well matched. We found no statistically significant differences between children in the three conditions in terms of age (in months), $F(2, 96) = 0.53$, $p = .59$; gender, χ^2 (2, $N = 104$) = 3.62, $p = .16$; ethnicity, χ^2 (8, $N = 104$) = 7.02, $p = .53$, or enrollment in the school with the greatest percentage of children qualifying for free or reduced lunch, χ^2 (2, $N = 104$) = 0.64, $p = 0.73$. We did, however, find a difference among conditions in terms of the grade level of the children, equivalent values $M = 2.81$ ($SD = 0.52$), iterative $M = 2.45$ ($SD = 0.52$), no extra practice $M = 2.81$ ($SD = 0.49$), $F(2, 101) = 8.07$, $p = .001$. Previous studies have been mixed in terms of whether they have found an association between grade level and performance solving math equivalence problems. Some have shown no difference in performance between children in grades 2-5 (Alibali, 1999; Falkner et al., 1999; NCISLA, 2000), whereas others have shown that second graders are actually more likely than third graders to solve math equivalence problems correctly (McNeil 2007, 2008). Thus, we tested and controlled for any effects of grade level in our analyses.

All problems in the practice sessions were single-digit addition problems with two addends, and all were presented in the traditional "operations on left side" format (i.e., $a + b =$ __). Thus, the practice sessions did not include any practice with math equivalence problems prior to the assessments. After receiving practice (or no input in the "no extra practice" condition), children completed posttest measures of understanding of math equivalence and computational fluency. Our goal was to examine if practicing simple addition facts grouped by equivalent values causes children to construct a better understanding of math equivalence.

Previous research has shown that simple exposure to math equivalence problems improves children's performance on the problems on later tests (Alibali & Meredith, 2009). Thus, it was essential that children not be exposed to our particular measures of interest prior to receiving an intervention and posttest. If children had been exposed to the measures beforehand, critics could argue that any improvements in understanding of math equivalence were due to the combination of exposure to math equivalence problems and organized arithmetic practice, rather than to the organized arithmetic practice itself.

Procedure

Figure 1 presents the general timing of the arithmetic practice sessions, assessment, and follow-up for children in each condition. Children in the two practice conditions participated individually in four sessions: three practice sessions—the last of which also included the posttest assessment—and one follow-up assessment occurring approximately two weeks after the third session. During the first three sessions, children practiced addition facts by playing games one-on-one with a tutor (all three sessions) and by answering flashcards (first two sessions). In between these sessions, children practiced by completing brief paper-and-pencil homework assignments. Overall, children in the practice conditions received approximately 100 minutes of practice prior to completing the assessments. Immediately following the third practice session, children were introduced to a new experimenter who administered a posttest to assess their understanding of math equivalence and computational fluency (see Assessment and Coding section). This new experimenter was used for the purposes of assessment, so he or she could be kept uninformed of children's practice condition. Children in the no extra practice condition completed the posttest assessment at the same point in time as children in the practice conditions, but they had not received any practice sessions yet. After completing the assessment, children in this no extra practice condition were randomly assigned to the equivalent values condition or to the iterative condition, and immediately began to receive practice sessions analogous to the two practice conditions in the main experiment. Finally, children in all conditions completed a follow-up assessment approximately two weeks after their final practice session. Thus, the delay between the final practice session and the follow-up assessment was the same for all children, regardless of condition. This “waitlist control” method enabled all children to benefit from the experiment by receiving organized practice with addition, and it allowed us to include all children in our analysis of the follow-up assessment. Most importantly, it also enabled us to perform a miniature pretest-intervention-posttest design with the subset of participants who were originally assigned to the no extra practice condition.

Practice sessions. The sessions were designed to help children practice solving single-digit addition facts with two addends (e.g., 9 + 8 = 17, 8 + 6 = 14). Children received practice via three main types of activities: (a) two-player games involving cards or dice, (b) flashcards, and (c) a computer game. One example of a two-player game is called the “Sorting Game”. In this game, the child and the experimenter each had a set of cards with a single-digit addition fact printed on each in the format $a + b =$ __. The goal of this game was to be the first person to organize his or her set of cards into two predetermined piles: “Tom’s pile” and “Mary’s pile.” Children were told that Tom and Mary each like a specific number. In the equivalent values condition, this number was the sum. In the iterative condition, this number was one of the addends. After sorting all of their cards into the two piles, children in the iterative condition were asked to solve each problem to ensure that they would receive the same amount of practice with each fact as children in the equivalent values condition.

Another example of a two-player game is called “Smack it!” (McNeil et al., 2011). In this game, the child and tutor each used a swatter with a suction cup on the end. At the beginning of each round, four solved addition problems were placed face-up on the table. Three of the problems were related to each other (e.g., in the equivalent values condition, the sum was always the same; in the iterative condition, one of the addends was always the same), and one of the problems was different. The goal was to be the first player to “smack” the addition problem that was different from the other three problems in the set. Children were then asked to explain why the problem was different from the other three. Children played a total of ten rounds. Children also played other two-player games that were similar in content and scope. The specific games played were the same in both practice conditions, and the games were rigged so all children practiced the same problems. The only difference between conditions was the manner in which problems were organized during the practice (by equivalent values or iteratively).

Children also practiced flashcards during the sessions. Before completing the flashcards, children received a brief demonstration on how to solve the flashcards. The tutor first presented a simple flashcard (e.g., “1 + 1 = __”). Next, the tutor read the problem aloud (e.g., “one plus one is equal to blank”). Finally, the tutor offered the correct number (e.g., “It’s two!”). When solving the flashcards, children were instructed to read each problem aloud before stating the number that should go in the blank. Thus, this flashcard practice was not designed to be a high-pressure task, as children were given ample time to read each problem aloud and to calculate their solution. Children in both practice conditions practiced all 81 addition facts (1 + 1 = __ through 9 + 9 = __) over the course of the first two practice sessions, but the flashcards were grouped into different sets based on condition. In the equivalent values condition, each set included all of the addend pairs that summed to the same value (e.g., 1 + 6 = __, 2 + 5 = __, 3 + 4 = __, 4 + 3 =

__, 5 + 2 = __, and 6 + 1 = __). In the iterative condition, each set included all of the addend pairs that started with the same first addend (e.g., 6 + 1 = __, 6 + 2 = __, 6 + 3 = __, 6 + 4 = __, 6 + 5 = __, 6 + 6 = __, 6 + 7 = __, 6 + 8 = __, and 6 + 9 = __). Children read and solved each flashcard in a set one at a time. After each flashcard was solved and feedback was given, the experimenter placed the solved flashcard face-up on the table. Once children finished solving all flashcards in the given set, the experimenter directed their attention to the flashcards on the table and asked them if they saw anything in common across all the cards. After children gave a reasonable response, the experimenter gave positive feedback and restated the commonality (e.g., “Good job! Each math problem is equal to seven!” or “Good job! Each math problem is six plus a number!”). Finally, the experimenter removed the flashcards from the table and moved on to the next set. Children completed a total of 45 flashcards during the first practice session and 36 flashcards during the second practice session.

Children also played a computer game in each of the three practice sessions. The game was a modified version of *Snakey Math* from Curry K. Software. The child and the tutor each controlled one of four snakes on the screen (the other two were computer controlled). Snakes could move up, down, left, and right. At the beginning of each round, a problem (e.g., 3 + 4 = __) was presented at the bottom of the screen and four numbers appeared in random locations on the screen. The goal was to be the first snake to eat the number that should go in the blank. Again, the only difference between conditions was the manner in which problems were organized during the game (by equivalent values or iteratively).

Homework. Children were asked to continue their practice in between practice sessions through brief homework assignments. These assignments were paper-and-pencil worksheets designed to take approximately 15 minutes to complete. All problems on the worksheets were single-digit addition problems with two addends and all were presented in the traditional "operations on left side" format (i.e., $a + b = c$). Thus, in line with the practice sessions, the worksheets did not include any math equivalence problems. Children were told that they would receive a sticker for completing the worksheets and turning them in when they returned for the next session. When a child turned in his or her completed worksheet, the tutor asked the child to read each problem aloud (along with the number written in the blank), and the tutor corrected any errors. Errors were rare; they occurred only on 1% of all problems solved by all children across both worksheets. To ensure that all children would be able to participate in this "read aloud" activity, tutors kept correctly completed worksheets on hand. Thus, children who had not turned in their own completed worksheet participated in the "read aloud" activity by using a worksheet that had been correctly completed for them. Most children turned in their own completed worksheets (97% completion rate from the first to second session, and 86% from the second to third session), and many who did not insisted that they had done the homework, but forgot to bring it with them. Every child returned at least one worksheet over the course of the intervention, and the number of children who failed to turn in at least one of the worksheets did not differ across the two practice conditions, χ^2 $(1, N = 70) = 1.34$, $p = .25$.

Assessments and coding

Posttest assessment of understanding of math equivalence. Children completed a three-component measure of understanding of math equivalence consisting of: (a) equation solving, (b) equation encoding, and (c) defining the equal sign. This measure was similar to measures used in many previous studies (e.g., McNeil & Alibali, 2005b; Hattikudur & Alibali, 2010; Matthews & Rittle-Johnson, 2009; McNeil et al., 2010). According to McNeil and Alibali (2005b), these three components tap a system of three distinct, but theoretically related constructs involved in children's understanding of math equivalence. We established inter-rater reliability on each component by having a second coder code the responses of 20% of the children.

To assess *equation solving*, children were videotaped as they solved and explained four math equivalence problems (1 + 5 = __ + 2, 7 + 2 + 4 = __ + 4, 2 + 7 = 6 + __, 3 + 5 + 6 = 3 + __). An experimenter placed each equation on an easel and said, "Try to solve the problem as best as you can, and then write the number that goes in the blank." After children wrote a number in the blank, the experimenter said, "Can you tell me how you got *x*" (*x* denotes the given answer) (cf. Alibali, 1999; Perry, 1991; Rittle-Johnson & Alibali, 1999; Siegler, 2002). Children's equation-solving strategies were coded as correct or incorrect based on a system used in previous research (e.g., McNeil & Alibali, 2004; Perry et al., 1988). Correctness was determined based on the solutions children wrote in the blank (e.g., for the problem 3 + 5 + 6 = 3 + __, a solution of 17 indicated an incorrect "Add All" strategy and a solution of 11 indicated a correct strategy). As in prior work, solutions were coded as reflecting a particular strategy as long as they were within ±1 of the solution that would be achieved with that particular strategy. Inter-rater reliability was high; agreement between coders on whether or not a given strategy was correct was 99%.

To assess *equation encoding*, children were asked to reconstruct four math equivalence problems (4 + 5 = 3 + __, 7 + 1 = __ + 6, 2 + 3 + 6 = 2 + __, 3 + 5 + 4 = __ + 4) after viewing each for 5 seconds (cf. Chase & Simon, 1973; Siegler, 1976). Children's encoding performance was coded as conceptually correct or conceptually incorrect based on a system used in previous research (e.g., McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999). Children were given credit for a conceptually correct reconstruction if they reconstructed the structure of the equation correctly. Thus, mistakes involving only the specific numbers or arrangements of the numbers (e.g., writing "7 + 1 = __ + 5" or "7 + 1 = 6 + __" for the equation 7 + 1 = __ + 6) were not counted as conceptual errors. Some common errors included converting the problem to a traditional addition problem (e.g., reconstructing 4 + 5 = 3 + __ as "4 + 5 + 3 = __"), omitting the plus on the right side of the equal sign (e.g., reconstructing 4 + 5 = 3 + __ as "4 + 5 = 3 __"), omitting the equal sign (e.g., reconstructing 7 + 1 = __ + 6 as "7 + 1 __ + 6." Agreement between coders on whether or not a given reconstruction was correct was 99%.

To assess *defining the equal sign*, children were videotaped as they responded to a set of questions about the equal sign. The experimenter pointed to an equal sign presented alone and asked: (1) "What is the name of this math symbol?" (2) "What does this math symbol mean?" and (3) "Can it mean anything else?" (cf. Behr et al., 1980; Baroody & Ginsburg, 1983; Knuth et al., 2006). Children's responses were categorized according to a system used in previous research (e.g., Knuth et al., 2006; McNeil & Alibali, 2005a, b). We were specifically interested in whether or not children defined the equal sign relationally as a symbol of math equivalence (e.g., "two amounts are the same"). Defining the equal sign relationally has had good concurrent validity in previous research (e.g., Knuth et al., 2006). Agreement between coders on whether or not a given definition was relational was 100%.

Follow-up assessment. All children also participated in a brief assessment of their equation-solving performance approximately two weeks after their third practice session. Children solved the same four math equivalence problems that they had solved during the original, posttest assessment. However, children were provided with brief tutelage and feedback at follow-up. When the tutor presented each equation, he or she instructed the child as follows: "You need to figure out what number goes in the blank (point to blank) to make this side of the equal sign (gesture to the left-hand side of the problem) the same amount as (point to the equal sign) this side of the equal sign (gesture to the right-hand side of the problem).” This brief tutelage was based on interventions used in previous studies (e.g., Alibali, 1999; McNeil & Alibali, 2005b; Perry, 1991; Rittle-Johnson & Alibali, 1999) that have been shown to help some, but not all, children solve math equivalence problems correctly, reconstruct the problems correctly after viewing them briefly, and define the equal sign relationally. If the child provided the correct number, the tutor gave positive feedback such as “good job” or “that’s right” and then moved on to the next problem. However, if the child provided an incorrect number, the tutor instructed the child as follows: “No, that’s not the number that goes in the blank. The correct number is *x* because *a* plus *b* is equal to *x* plus *y*” (Note: the actual numbers in the problem are used in place of *a*, *b*, *x*, and *y*). The purpose of this follow-up session was to examine potential enduring effects of the practice conditions on children’s openness to learn from brief instruction on math equivalence.

Measures of computational fluency. In addition to the measure of understanding of math equivalence, children also completed two measures of computational fluency. In the first measure, children solved a set of simple addition problems that were presented one at a time in the center of a computer screen (cf. Geary, Bow-Thomas, Liu, & Siegler, 1996; LeFevre, Sadesky, & Bisanz, 1996; Siegler, 1988). For each trial, a fixation point appeared in the center of the screen followed by the prompts “ready,” “set,” and “go,” displayed for 1 second each. The two addends were then presented on the screen (e.g., 9 + 8), and they remained on the screen until the child said his or her answer aloud. Accuracy and reaction time were recorded. The second measure was the Math Computation section of Level 8 of the Iowa Test of Basic Skills, which is a standardized, timed measure of arithmetic computation that yields a percentile rank.

Results

Children's understanding of math equivalence was poor overall. This finding is consistent with prior studies (e.g., Alibali, 1999; McNeil, 2008; Perry et al., 1988), and it was not surprising, given that all practice problems in the experiment were presented in the traditional $a + b = c$ format (see McNeil et al. 2011). Table 1 presents performance on each component of the math equivalence measure by condition. As shown in the table, children in the equivalent values condition performed best across all three components of the math equivalence measure. We analyzed the effects of practice condition on each of the three components of understanding of math equivalence and on the follow-up assessment.

Effects of practice condition on children's equation solving performance

Equation solving performance was poor overall (M = 0.67 [out of 4], SD = 1.39) and not normally distributed, with 79 of 104 children (76%) solving zero equations correctly. The distribution was bimodal with most children solving zero (76%) or all four (14%) equations correctly. Bimodal distributions of performance on math equivalence problems are common (e.g., McNeil & Alibali, 2004) and expected because it takes a dramatic shift in understanding just to solve one equation correctly. Thus, we classified children into two groups: those who solved zero equations correctly (N = 79) and those who solved at least one equation correctly (N = 25). We then used binomial logistic regression to predict the log of the odds of solving at least one equation correctly (see Agresti, 1996). Grade level was not a significant predictor of performance, $\hat{\beta} = 0.59$, $z = 1.34$, $Wald\,(1, N = 104) = 1.80$, $p = .18$, so we did not include it in the model. However, conclusions were the same when we included it.

To represent the three levels of practice condition, we used the set of two orthogonal Helmert contrast codes that mapped onto our a priori hypotheses: (1) equivalent values condition versus the two "control" conditions, and (2) iterative condition versus the no extra practice condition. As predicted, children in the equivalent values condition were more likely than children in the other two conditions to solve at least one equation correctly (13 of 36 [36%] versus 12 of 68 [18%]), $\hat{\beta} = 0.97$, $z = 2.06$, *Wald* $(1, N = 104) = 4.25$, $p = .04$. The effect size was medium; the model estimates that the odds of solving at least one equation correctly are more than 2.6 times higher after participating in the equivalent values condition than after participating in one of the control conditions. Children in the iterative and no extra practice conditions were equally likely to solve at least one equation correctly (6 of 34 [18%] in both conditions). Thus, as predicted, practice organized around equivalent values produced the best chance of solving a math equivalence problem correctly. A separate pairwise comparison of the equivalent values and iterative conditions was marginal, but consistent with the earlier analysis ($p = .08$).

We also examined strategy use in each of the three conditions. Table 2 presents the percentage of equations solved with each strategy in each condition. Two aspects of the table warrant attention. First, the equivalent values condition had the greatest percentage of equations solved with a correct strategy (22.2% vs. 13.2% for iterative and 14.7% for no extra practice). Second, the equivalent values condition had the lowest percentage of equations solved using either the Add All or Add to Equal Sign strategies (59.1% vs. 68.4% for iterative and 63.9% for no extra practice). These two strategies are the ones most commonly used by children who have not received any instruction on math equivalence (McNeil, 2005). They are both incorrect generalizations from children's experience with traditional arithmetic, as children are simply adding up all the numbers in the equation or adding up all the numbers to the left of the equal sign, rather than considering the implications of having addends on both sides of the equation.

Effects of practice condition on children's encoding performance

Encoding performance was poor overall (M = 1.41 [out of 4], SD = 1.45) and not normally distributed, with 38 of 104 children (37%) reconstructing zero equations correctly. Thus, we again categorized children into two groups: those who reconstructed zero equations correctly (N = 38) and those who reconstructed at least one equation correctly (N = 66) and used binomial logistic regression. Grade level was not a significant predictor of performance, $\hat{\beta}$ = 0.43, z = 1.13, *Wald* (1, N = 104) = 1.28, p = .26, so we did not include it in the model. However, conclusions were the same when we included it.

As predicted, children in the equivalent values condition were more likely than children in the other two conditions to reconstruct at least one equation correctly (28 of 36 [78%] versus 38 of 68 [56%]), $\hat{\beta}$ = 1.02, z = 2.17, *Wald* (1, N = 104) = 4.68, p = .03. The effect size was medium; the model estimates that the odds of reconstructing at least one equation correctly are over 2.7 times higher after participating in the equivalent values condition than after participating in one of the control conditions. Children in the iterative condition were not statistically different from children in the no extra practice condition, (20 of 34 [59%] versus 18 of 34 [53%]), $\hat{\beta}$ = -0.24, z = -0.49, *Wald* (1, N = 68) = 0.24, p = .63. Thus, as predicted, practice organized around equivalent values produced the best chance of encoding a math equivalence problem correctly. A separate pairwise comparison of the equivalent values and iterative conditions led to similar conclusions (p = .09).

Effects of practice condition on children's equal sign definition

Equal sign defining was poor, with only 7 of 104 children (7%) defining the equal sign relationally. Grade level was a significant predictor of performance, $\hat{\beta}$ = 2.12, z = 1.97, *Wald* (1, N = 104) = 3.88, p = .05, so we included it in the model. Controlling for grade level, children in the equivalent values condition were not significantly more likely than children in the other two conditions to define the equal sign relationally (4 of 36 [11%] versus 3 of 68 [4%]), $\hat{\beta} = 0.84$, $z = 1.00$, *Wald* $(1, N = 104) = 1.00$, $p = .32$. Children in the iterative condition were not statistically different from children in the no extra practice condition, (2 of 34 [6%] versus 1 of 34 [3%]), $\hat{\beta} = 0.03$, $z = 0.02$, *Wald* $(1, N = 68) = 0.001$, $p = .98$.

A marginal advantage of the equivalent values condition emerged when we were slightly more lenient and counted "is equal to" and "equal to" as relational definitions. We typically do not give credit for these underspecified definitions; however, some researchers have been far more lenient in their coding (e.g., Rittle-Johnson & Alibali, 1999), and other researchers have not specified how they treat such responses (e.g., Hattikudur & Alibali, 2010). When applying the more lenient criterion, 12 of 104 children (12%) defined the equal sign relationally. Even after controlling for grade level, children in the equivalent values condition were marginally more likely than children in the other two conditions to do so (7 of 36 [19%] versus 5 of 68 [7%]), $\hat{\beta} = 1.22$, $z = 1.69$, *Wald* $(1, N = 104) = 2.85$, $p = .09$. The effect size was large; the model estimates that the odds of providing a relational definition are 3.4 times higher after participating in the equivalent values condition than after participating in one of the control conditions.

Effects of practice condition on computational fluency

Because there are sometimes trade-offs between conceptual understanding and computational fluency, it was important to test if the gains in understanding of math equivalence for the equivalent values condition (shown above) were accompanied by decrements in computational fluency when compared to the iterative condition. We did not find any evidence of such trade-offs in this sample. Children in the equivalent values condition had a similar average Normal Curve Equivalent ($M = 50.48$, $SD = 13.49$) to children in the iterative ($M =$

51.04, $SD = 17.69$) condition on Level 8 of the Math Computation section of the Iowa Tests of Basic Skills, $F(1, 68) = 0.023$, $p = .88$. They also performed similarly when solving single-digit addition facts: accuracy (equivalent values $M = 13.36$, $SD = 1.10$; iterative $M = 12.88$, $SD = 1.83$), $F(1, 68) = 1.77$, $p = .19$, average reaction time (equivalent values $M = 5.08$ s, $SD = 1.93$; iterative $M = 5.91$ s, $SD = 2.71$), $F(1, 68) = 2.22$, $p = .14$), and number of problems on which retrieval was used (equivalent values $M = 2.72$, $SD = 3.04$; iterative $M = 3.44$, $SD = 3.48$, $F(1, 68) = 0.89$, $p = .36$. This is not surprising given that children in both practice conditions received organized practiced with the same set of addition facts over the course of the intervention (only the specific organization of the facts differed).

We also found tentative evidence that our two practice conditions facilitated children's computational fluency. Specifically, children in the no extra practice condition had a lower average Normal Curve Equivalent ($M = 38.87$, $SD = 14.23$) than children in the two practice conditions, $F(1, 101) = 13.98$, $p < .001$. They also had lower use of retrieval when solving single-digit addition facts ($M = 1.76$, $SD = 2.80$) than children in the two practice conditions, $F(1, 101) = 4.08$, $p = .046$, but they had similar accuracy ($M = 13.09$, $SD = 1.14$) and reaction time ($M = 5.71$ s, $SD = 2.61$).

Effects of practice condition on follow-up performance

We lost four children to attrition (3 equivalent values and 1 iterative) at follow-up. Recall that children received brief tutelage prior to completing the follow-up assessment. They also received feedback after solving each equation. Also recall that there were only two conditions to compare on the follow-up assessment: equivalent values and iterative. This is because the no extra practice condition was a "waitlist control" condition in which children were randomly assigned to one of the two practice conditions after taking the posttest. All children regardless of condition participated in the follow-up assessment approximately two weeks after their third practice session. Grade level was not a significant predictor of performance, $\hat{\beta} = 0.41$, $z = 1.09$, *Wald* $(1, N = 100) = 1.80$, $p = .28$, so we did not include it in the model. However, conclusions were the same when we included it.

Overall, performance was better than it was on the posttest, but it was still low ($M = 1.65$ [out of 4], $SD = 1.62$) and not normally distributed, with 40 of 100 (40%) of children solving zero equations correctly. We again categorized children into two groups: those who solved zero equations correctly ($N = 40$) and those who solved at least one equation correctly ($N = 60$) and used binomial logistic regression. As predicted, children in the equivalent values condition were more likely than children in the iterative condition to solve at least one equation correctly (35 of 50 [70%] versus 25 of 50 [50%]), $\hat{\beta} = 0.85$, $z = 2.02$, *Wald* $(1, N = 100) = 4.10$, $p = .04$. The effect size was medium; the model estimates that the odds of solving at least one equation correctly are more than 2.3 times higher after participating in the equivalent values condition than after participating in the iterative condition.

We also examined strategy use in the two conditions at follow-up. Table 3 presents the percentage of equations solved with each strategy in each condition. The two patterns that emerged at posttest remained at follow-up. First, the percentage of equations solved with a correct strategy was greater in the equivalent values condition (50%) than in the iterative condition (33%). Second, the percentage of equations solved using either the Add All or Add to Equal Sign strategies was lower in the equivalent values condition (27.5%) than in the iterative condition (38.5%).

Supplemental analysis of children in the no extra practice condition

We separately analyzed the follow-up performance of children who had originally been assigned to the no extra practice condition (N = 34). These children took the "posttest" before they had been randomly assigned to one of the two practice conditions, so we knew how well they performed on the math equivalence measures prior to receiving the intervention. Thus, this subset of children participated in a randomized experiment with a pretest, intervention, and follow up. There were no significant differences between children in the two conditions in terms of demographic variables or scores at pretest. Table 4 presents pre-intervention scores on each component of the math equivalence measure by condition along with scores on the follow-up assessment. We used binomial logistic regression to predict the log of the odds of solving at least one equation correctly on the follow-up. We included pre-intervention scores on the encoding, equation solving, and equal sign defining measures as control variables, but conclusions are the same if they are not included. As predicted, even after controlling for pre-intervention scores, children who ended up receiving the equivalent values practice were more likely than children who ended up receiving the iterative practice to solve at least one equation correctly at follow-up (13 of 14 [93%] versus 6 of 16 [38%]), $\hat{\beta}$ = 3.36, z = 2.54, *Wald* (1, N = 30) = 6.44, p = .01. The effect size was large; the model estimates that the odds of solving at least one equation correctly are more than 28 times higher after receiving the equivalent values practice than after participating in the iterative practice. Note that the particularly high performance children in the equivalent values condition in this supplemental analysis reinforces our concerns about pretest sensitization. These children received the same amount of practice as children in the original equivalent values condition, and they also had same length of delay between their final practice session and follow-up assessment. Nonetheless, they performed significantly better than children in the original equivalent values condition on the follow-up (13 of 14 [93%] versus 22 of 36

[61%]), χ^2 (1, $N = 50$) = 4.84, p =. 03. This result suggests that children may reap more benefits from arithmetic practice that is organized by equivalent values if they have been exposed to math equivalence problems prior to the practice than if they have not been exposed prior to the practice.

Discussion

The present experiment provides the first evidence that the organization of children's knowledge of arithmetic facts may be an important source of individual differences in children's understanding of math equivalence. We randomly assigned children to practice arithmetic in one of three conditions. Children who practiced problems that were grouped by equivalent values came away with a better understanding of math equivalence than children who practiced problems grouped in an iterative fashion or children who received no extra practice. Importantly, the gains in understanding of math equivalence were not accompanied by any detectable difference in computational fluency. These results provide evidence for a causal link between a conceptual organization of arithmetic facts and a better understanding of math equivalence. It is noteworthy that this link exists even when the format of the addition facts takes the standard "operations on left side" format that has been implicated in children's misconceptions (e.g., McNeil et al., 2011).

Potential Mechanisms

Our evidence indicates that the external organization of arithmetic practice affects children's understanding of math equivalence. However, the mechanisms underlying this effect are less clear. How is it that conceptually based organization of arithmetic facts helps children to form a better understanding of math equivalence? Although it is only speculation at this point, we have proposed three possibilities. First, arithmetic practice organized around equivalent values may help children extract conceptual knowledge about both number and math equivalence because it exposes children to multiple, consecutive examples of the same value that vary in surface details. For instance, the group of facts that sum to 9 (e.g. 3 + 6, 4 + 5, etc.), vary in appearances but are structurally equivalent in that they sum to 9. According to Gelman and Williams (1998), this type of exposure to structurally redundant examples in a domain increases the chances that a learner's mental structures will be compatible with the underlying structure of the domain. More specifically, this organization may increase children's chances of inducing arithmetic principles, such as the commutative property, from their experience with arithmetic. As Carpenter et al.'s work (e.g., Falkner et al., 1999; Jacobs et al., 2009; NCISLA, 2000) has convincingly demonstrated, a key strategy for helping children construct a good understanding of mathematical equivalence is to teach them to notice and understand the structure and properties of arithmetic.

Second, arithmetic practice organized around equivalent values may help children process math equivalence problems more efficiently because it leads equivalent addend pairs to become associated. When facts with equivalent values are practiced in succession, multiple terms—including the sum (e.g., 9) *and* addend pairs sharing that sum (e.g., 3 + 6, 2 + 7, etc.)—come to be increasingly associated with one another. Subsequently, when faced with an equation that includes one of the addend pairs, equivalent addend pairs may be primed. Once concurrently activated, these equivalent combinations have the potential to influence encoding of, strategy selection on, and/or conceptualization of equations involving these facts via top-down processes (cf. Bruner, 1957; McGilly & Siegler, 1990; Rumelhart, 1980).

Third, practice with problems organized by equivalent sums may enable children to notice that addend pairs can be equivalent to one another based on transitive relations (e.g., if 3 +

4 = 7 and 5 + 2 = 7, then 3 + 4 = 5 + 2). This particular insight—that an addend pair can be equivalent not only to a single number, but also to other addend pairs—is a key milestone in the construction of an understanding of math equivalence (Behr et. al., 1980; Kieran, 1981). When children are faced with the fact that two addend pairs can be equivalent to one another, it challenges their misconceptions about the way arithmetic works and facilitates relational thinking about numbers (Falkner et al., 1999; Jacobs et al., 2007; NCISLA, 2000).

Practical Tools for Teaching Math equivalence

In addition to theoretical implications for mechanisms of knowledge construction, our results have important practical implications. We have identified the organization of arithmetic practice as a malleable factor that can be changed to improve children's understanding of math equivalence. Our findings support the idea that seemingly minor differences in children's early input (i.e. the organization of arithmetic facts practice) can play a central role in shaping and constraining the path of development of children's understanding of fundamental concepts (see also Klibanoff, et al. 2006; McNeil et al., 2011; Ramani & Siegler, 2008). A major strength of our findings is the practicality of the intervention—it requires neither extensive teacher training, nor months of explicit conceptual instruction for children, yet it is still productive in preparing children to learn the concept of math equivalence. However, it must be acknowledged that our single intervention by no means led children to form a mastery level understanding of math equivalence. Performance remained relatively poor overall, and children's definitions of the equal sign did not benefit as much as their equation solving and encoding performance. A series of future studies has been designed to test whether using this intervention combined with other instructional interventions can lead children to form a more robust and comprehensive understanding of math equivalence.

Limitations and Future Directions

Although our results support the hypothesis that organizing arithmetic facts based on equivalent total values is a key factor supporting the formation of an understanding of math equivalence, there are at least three limitations. First, our use of tightly controlled, scripted activities during the intervention raises concerns regarding the generalization of our results to different learning environments. Nevertheless, the simplicity of our manipulation, which merely requires using a specific organization of the addition facts already practiced routinely in schools and homes throughout the U.S., increases the likelihood that our findings will generalize to less-controlled learning environments such as formal classrooms or parent-child interactions.

Second, we did not gather pretest measures of understanding of math equivalence on the full sample because it was important to avoid pretest sensitization. Such sensitization was a real concern because previous research has shown that simple exposure to math equivalence problems improves children's performance on later tests of math equivalence (Alibali & Meredith, 2009). Thus, we used a posttest-only design and randomly assigned individual children to conditions. Random assignment was successful in creating groups that were similar in terms of most demographics and context variables, but fell short in terms of equating groups by grade level. Thus, we cannot rule out the possibility that there were some other grade-related differences in mathematical understanding between groups in the main experiment at pretest. Fortunately, grade level (2-6) has not consistently been associated with knowledge of math equivalence (Alibali, 1999; Falkner et al., 1999; NCISLA, 2000), and controlling for it in the present analyses did not affect conclusions.

Importantly, we supplemented our main experiment with a manipulation of our waitlist control group to serve as a check on the conclusions of our larger design. With this method, we randomly assigned children who were originally assigned to the no extra practice condition to one of the practice conditions after they had completed the posttest assessment. Thus, the nominal posttest effectively served as a pretest for this subset of children, and the nominal follow-up assessment served as a delayed posttest. Again, there were no pre-intervention differences in demographics or understanding of math equivalence for this subset of children, and we still found that post-intervention performance for the equivalent values condition was better than performance for the iterative condition. The results from this waitlist control manipulation were even stronger than the findings of our main experiment and help mitigate potential concerns with our larger posttest-only design.

Third, we do not wish to imply that conceptually organized arithmetic practice is strictly necessary or sufficient for the formation of a nuanced understanding of math equivalence. Our goal was to test a theoretical claim about one potential source of individual differences in children's understanding of math equivalence, not to design an intervention to get children to mastery-level understanding. However, given that the ultimate goal is to leverage all of the recent theoretical advances to design a comprehensive intervention, the present intervention may be useful in combination with other effective methods for cultivating an understanding of math equivalence. Some of these other methods include: (1) practicing arithmetic problems in nontraditional formats such as __ = 9 + 8 or 17 = 9 + __ (Baroody & Ginsburg, 1983; Denmark, Barco, & Voran, 1976; MacGregor & Stacey, 1999; McNeil et al., 2011; Seo & Ginsburg, 2003; Weaver, 1973), (2) learning about the equal sign in non-arithmetic contexts, such as 28 = <u>28</u> or 1 foot = <u>12</u> inches (McNeil, 2008; see also Li et al. 2008 for comparison to educational practices used in China), (3) solving equations presented with concrete materials (e.g., wooden blocks, stickers) before "fading" to more abstract symbols (Sherman & Bisanz, 2009; Fyfe & McNeil, 2009), (4) comparing and explaining different types of problems and strategies (e.g., Falkner, et al., 1999; Hattikudur & Alibali, 2010; Jacobs et al., 2007; Rittle-Johnson, 2006; Siegler, 2002).

Conclusions

The present study adds to a growing literature suggesting that relatively minor, practical modifications to children's early learning environments can affect their understanding of foundational mathematical concepts (e.g., McNeil et al., 2011; Ramani & Siegler, 2008). Results reinforce the idea that children benefit when the structure of early input is modified to map onto the underlying structure of the conceptual domain of interest. More specifically, results indicate a causal link between a conceptual organization of arithmetic facts practice and performance solving and encoding equations with operations on both sides of the equal sign. That is, simply organizing arithmetic facts practice by equivalent total values helps children better understand math equivalence. Such findings not only contribute to our general understanding of how children construct knowledge, but also offer specific methods for facilitating the learning of an important mathematics concept that is necessary for success in algebra.

References

Adelman, C. (2006). The toolbox revisited: paths to degree completion from high school through college. *US Department of Education*.

Agresti, A. (1996). *An introduction to categorical data analysis.* New York: Wiley.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145. doi:10.1037/0012-1649.35.1.127

Alibali, M. W., & Meredith, L. (2009). Sources of change in children's encoding of mathematics problems. In B. Rittle-Johnson (Organizer), *Understanding knowledge change: Investigations on how children learn mathematics and literacy skills.* Symposium presented at the Biennial Meeting of the Cognitive Development Society (CDS), San Antonio, TX.

Alibali, M. W., Phillips, K. M. O., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development, 24*, 89-101. doi:10.1016/j.cogdev.2008.12.005

Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal, 84*, 199-212. doi:10.1086/461356

Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equal sign. *Mathematics Teaching, 92*, 13-15.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education, 36*, 412-446.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Revue of Psychology, 53,* 371-399. doi: 10.1146/annurev.psych.53.100901.135233

Bruner, J. S. (1957). On perceptual readiness. *Psychonomic Review, 2*, 123-152.

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

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian approach. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81. doi:10.1016/0010-0285(73)90004-2

Chi, M.T.H., & Ceci, S.J. (1987). Content knowledge: Its role, representation, and restructuring in memory development. In L. Lipsitt (Ed.), *Advances in Child Development and Behavior* (Vol. 20, pp. 91-143). New York: Academic Press. doi:10.1016/S0065-2407(08)60401-2

Cobb, P. (1987). An investigation of young children's academic arithmetic contexts. *Educational Studies in Mathematics*, *18*, 109-124. doi:10.1007/BF00314722

Denmark, T., Barco, E., & Voran, J. (1976). *Final report: A teaching experiment on equality. PMDC Technical Report No. 6.* Tallahassee, FL: Florida State University.

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, December, 232-236.

Fyfe, E. R., & McNeil, N. M. (2009). Benefits of "concreteness fading" for children with low knowledge of math equivalence. Poster presented at the Biennial Meeting of the

Cognitive Development Society (CDS), San Antonio, TX.

Geary, D. C., Bow Thomas, C. C., Liu, F., & Siegler, R. S. (1996) Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67*, 2022-2044. doi:10.2307/1131607

Gelman, R., and Williams, E. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In D. Kuhn and R. Siegler, (Eds.). Cognition, perception and language. Vol. 2. Handbook of Child Psychology (Fifth Ed). (pp. 575-630). W. Damon, Editor-in-Chief; New York: John Wiley and Sons.

Hattikudur, S., & Alibali, M. W. (2010). Learning about the equal sign: Does comparing with inequality symbols help? *Journal of Experimental Child Psychology*, *107*, 15-30. doi:10.1016/j.jecp.2010.03.004

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education, 38,* 258-288.

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28*, 644-653. doi:10.1037/0012-1649.28.4.644

Kieran, C. (1980). The interpretation of the equal sign: Symbol for an equivalent relation vs. an operator symbol. In *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education* (pp. 163-169). Berkeley: University of California, Lawrence Hall of Science.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in*

Mathematics, 12, 317-326. doi:10.1007/BF00311062

Kieran, C. (1992). The learning and teaching of school algebra. *Handbook of research on mathematics teaching and learning*, 390-419.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006) Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology, 42,* 59-69. doi:10.1037/0012-1649.42.1.59

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*, 297-312.

LeFevre, J.A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 216-230. doi:10.1037/0278-7393.22.1.216

Li, X., Ding, M., Capraro, M. M., & Capraro, R. M. (2008). Sources of differences in children's understandings of mathematical equality: Comparative analysis of teacher guides and student texts in China and the United States. *Cognition and Instruction, 26*, 195-217. doi:10.1080/07370000801980845

Lindberg, M. A. (1980). Is knowledge base development a necessary and sufficient condition for memory development? *Journal of Experimental Child Psychology, 30*, 401-410. doi:10.1016/0022-0965(80)90046-6

MacGregor, M., & Stacey, K. (1999). A flying start to algebra. *Teaching Children Mathematics, 6,* 78-85.

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.

McGilly, K. & Siegler, R. S. (1990). The influence of encoding and strategic knowledge on children's choices among serial-recall strategies. *Developmental Psychology, 26*, 931-941. doi: 10.1037/0012-1649.26.6.931

McNeil, N. M. (2005). *When the basics backfire: The change-resistance account of equation-learning difficulties.* Unpublished doctoral dissertation, University of Wisconsin, Madison.

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology, 43*, 687-695. doi:10.1037/0012-1649.43.3.687

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of math equivalence. *Child Development, 79*, 1524-1537. doi:10.1111/j.1467-8624.2008.01203.x

McNeil, N. M. (2009, June). *Modifying Arithmetic Practice to Promote Understanding of Mathematical Equivalence.* Invited talk presented at the Annual Research Conference of the Institute of Education Sciences (IES), Washington, DC. Retrievable from http://ies.ed.gov/director/conferences/09ies_conference/presentations.asp

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466. doi:10.1207/s15516709cog2803_7

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6*, 285-306. doi:10.1207/s15327647jcd0602_6

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899. doi:10.1111/j.1467-8624.2005.00884.x

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. S., Hattikudur, S., & Krill, D. E. (2006). Middle-school students' understanding of the equal sign: The books they read can't help. *Cognition and Instruction, 24,* 367-385. doi:10.1207/s1532690xci2403_3

McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shipley, H. (2011). Benefits of practicing 4 = 2 + 2: Nontraditional problem formats facilitate children's understanding of math equivalence. *Child Development, 82*(5), 1620-1633. doi: 10.1111/j.1467-8624.2011.01622.x.

McNeil, N. M., Rittle-Johnson, B., Hattikudur, S. & Petersen, L. A. (2010). Continuity in representation between children and adults: Arithmetic knowledge hinders undergraduates' algebraic problem solving. *Journal of Cognition and Development*, *11*(4), 437-457. doi:10.1080/15248372.2010.516421

Moses, R. P., & Cobb, C. E. (2001). *Radical equations: Math literacy and civil rights*. Beacon Press. doi:10.2307/2678804

National Center for Improving Student Learning and Achievement in Mathematics Education (2000). *Building a foundation for learning algebra in the elementary grades.* Madison, WI: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academy Press.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6,* 449-468. doi:10.1016/0885-2014(91)90049-J

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400. doi:10.1016/0885-2014(88)90021-4

Piaget, J. & Szeminska, A. (1995). The child's conception of number (C. Gattegno & F. M. Hodgson, Trans.). In H. E. Gruber & J. J. Voneche (Eds.), *The essential Piaget* (pp. 298-341). Northvale, NJ: Jason Aronson Inc. (Original work published 1941).

Powell, S. R., & Fuchs, L. S. (2010). Contribution of equal-sign instruction beyond word problem tutoring for third-grade students with mathematics difficulty. *Journal of Educational Psychology*, *102*(2), 381-394. doi:10.1037/a0018447

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79,* 375-394. doi:10.1111/j.1467-8624.2007.01131.x

Renwick, E. (1932). Children's misconceptions concerning the symbols for mathematical equality. *British Journal of Educational Psychology, 2*, 173-183.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77,* 1-15. doi:10.1111/j.1467-8624.2006.00852.x

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189. doi:10.1037/0022-0663.91.1.175

Rumelhart, D. E. (1980). Schemata: The building block of cognition. In R. Spiro, B. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hilsdale, NJ: Erlbaum.

Seo, K.-H., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence...": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 161-187). Mahwah, NJ: Lawrence Erlbaum Associates.

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and non-symbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology, 101,* 88-100. doi:10.1037/a0013156

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481-520. doi:10.1016/0010-0285(76)90016-5

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*, 833-851. doi:10.2307/1130252

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York: Cambridge University. doi:10.1017/CBO9780511489709.002

Steinberg, R. (1985). Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education*, 16, 337-355. doi:10.2307/749356

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1990). Algebra students' knowledge of equivalence of equations. *Journal for Research in Mathematics Education*, *22*, 112-121.

Tran, H., & Weinraub, M. (2006). Child care effects in context: Quality, stability, and multiplicity in nonmaternal child care arrangements during the first 15 months of life. *Developmental Psychology, 42,* 566-582. doi:10.1037/0012-1649.42.3.566

Weaver, J. F. (1973). The symmetric property of the equality relation and young children's ability to solve open addition and subtraction sentences. *Journal for Research in Mathematics Education, 4*, 45-56. doi:10.2307/749023

Table 1

Performance on Each Component of Understanding of Math Equivalence by Condition

Component & measure	Equivalent values	Iterative	No extra practice
Equation solving			
M (*SD*)	0.88 (1.48)	0.53 (1.30)	0.59 (1.37)
% at least one correct	36%	17%	18%
Equation encoding			
M (*SD*)	1.17 (1.00)	0.94 (1.04)	0.82 (1.14)
% at least one correct	69%	53%	44%
Defining the equal sign			
% who defined relationally	11%	3%	6%

Table 2

Percentage of Equations Solved With Each Strategy in Each Condition at Posttest

Strategy	Definition	Eq values	Iterative	No extra
Correct	Any strategy that makes both sides of the equation equal.	22.2%	13.2%	14.7%
Add All	Add all the numbers in the equation.	28.5%	30.9%	36.0%
Add to Equal Sign	Add all the numbers appearing to the left of the equal sign.	30.6%	37.5%	27.9%
Carry	Take one of the numbers appearing to left of the equal sign and write it in the blank.	6.3%	5.1%	8.8%
Other Incorrect	Incorrect strategies other than Add All, Add to Equal Sign, or Carry.	12.5%	13.2%	12.5%

Table 3

Percentage of Equations Solved With Each Strategy in Each Condition at Follow-up

Strategy	Eq values	Iterative
Correct	50.0%	33.0%
Add All	9.5%	16.0%
Add to Equal Sign	18.0%	22.5%
Carry	5.5%	7.0%
Other Incorrect	17.5%	21.5%

Table 4

Pre-intervention and Follow-up Performance by Condition for the Subset of Children Who Had Originally Been Assigned to the No Extra Practice Condition

Time point	Component & measure	Equivalent values	Iterative
Pre-intervention	Equation solving		
	M (*SD*)	0.71 (1.45)	0.47 (1.34)
	% at least one correct	24%	12%
	Equation encoding		
	M (*SD*)	1.06 (1.25)	0.59 (1.00)
	% at least one correct	53%	35%
	Defining the equal sign		
	% who defined relationally	0%	12%
Follow-up	Equation solving		
	M (*SD*)	2.71 (1.26)	1.13 (1.67)
	% at least one correct	93%	38%

Condition	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
Equivalent values	Practice	Practice	Practice then Assessment		Follow-up		
Iterative	Practice	Practice	Practice then Assessment		Follow-up		
No extra practice			Assessment then Practice	Practice	Practice		Follow-up

* Note. Children in all conditions completed practice worksheets for homework after their first and second practice sessions.

Figure 1. General timing of arithmetic practice sessions, assessment, and follow-up for children in each condition.