Cognitive Development 30 (2014) 30–46

Contents lists available at ScienceDirect

Cognitive Development

Organization matters: Mental organization of addition knowledge relates to understanding math equivalence in symbolic form☆

Dana L. Chesney[a,*], Nicole M. McNeil[a,**], Percival G. Matthews[a,1], Caroline E. Byrd[a], Lori A. Petersen[a,2], Mary C. Wheeler[a], Emily R. Fyfe[b], April E. Dunwiddie[a]

[a] *Department of Psychology, University of Notre Dame, United States*
[b] *Department of Psychology and Human Development, Vanderbilt University, United States*

ARTICLE INFO

Keywords:
Conceptual development
Mental organization
Mathematical understanding
Mathematical knowledge
Individual differences

ABSTRACT

Two quasi-experiments examined mental organization of addition knowledge as a potential source of individual differences in understanding math equivalence in symbolic form. We hypothesized that children who mentally organize addition knowledge around conceptually related groupings would have better understanding of math equivalence. In Quasi-experiment 1, we assessed 101 second and third grade students' mental organization of addition knowledge based on their use of decomposition strategies to solve addition problems (e.g., 3 + 4 = 3 + 3 + 1 = 6 + 1 = 7). In Quasi-experiment 2, we assessed 94 second grade students' mental organization based on their ability to generate a set of equations equal to a target value. In both quasi-experiments, children whose mental organization better reflected conceptually related groupings exhibited better understanding of math equivalence. Results thus support the

☆ This research was supported by grant R305B070297 from the U.S. Dept. of Education, Institute of Education Sciences. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Dept. of Education. Thanks to Martha Alibali and Heather Brletic-Shipley for their input and Krysten Williams and Jennifer Heil for their help with coding. This research would not have been possible without the support of many administrators, teachers, parents, and students in the greater South Bend, IN and Washington, DC areas.

* Corresponding author. Present address: Department of Psychology, The Ohio State University, 225 Psychology Building, 1835 Neil Avenue, Columbus, OH, 43210, United States. Tel.: +1 862 432 6539.
** Corresponding author at: Department of Psychology, University of Notre Dame, 118 Haggar Hall, Notre Dame, IN, 46556, United States. Tel.: +1 574 631 5678.
E-mail addresses: dlchesney@gmail.com (D.L. Chesney), nmcneil@nd.edu (N.M. McNeil).
[1] Present address: University of Wisconsin-Madison, United States.
[2] Present address: University of Chicago, United States.

http://dx.doi.org/10.1016/j.cogdev.2014.01.001

hypothesis that mental organization of addition knowledge into conceptually related groupings based on equivalent values may influence understanding of math equivalence in symbolic form.

1. Introduction

Although many developmental psychologists seek to identify commonalities in children's development, studying individual differences (Cronbach, 1957; Underwood, 1975) may provide insight into mechanisms of typical development (Hughes et al., 2005; Nelson, 1981) and inspire interventions that facilitate learning in reading (Blachman, Tangel, Ball, Black, & McGraw, 1999; Shaywitz et al., 2004) and mathematics (Booth & Siegler, 2006; Ramani & Siegler, 2008). Here, we examine a source of individual differences in children's understanding of math equivalence.

1.1. Mathematical equivalence

Mathematical equivalence, commonly symbolized by the equal sign (=), is the relation between two interchangeable quantities (Kieran, 1981). Understanding math equivalence in symbolic form not only involves understanding the meaning of the equal sign, but also encoding math equations in their entirety, correctly identifying an equation's two "sides," and noticing relations within equations (Jacobs, Franke, Carpenter, Levi, & Battey, 2007; Rittle-Johnson & Alibali, 1999). To be concise, we herein refer to this array of knowledge as "understanding of math equivalence", although we are specifically referring to understanding of math equivalence in symbolic form.

Understanding of math equivalence is critical to development of algebraic thinking (Falkner, Levi, & Carpenter, 1999; Kieran, 1992; Knuth, Stephens, McNeil, & Alibali, 2006). Unfortunately, most U.S. children have poor understanding of math equivalence (Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols, 1980; Falkner et al., 1999; McNeil, 2008; Perry, 1991). McNeil (2005) found that nearly 80% of U.S. 7–11-year-olds solve math equivalence problems—problems with operations on both sides of the equal sign (e.g., $6+3=4+\_$)—incorrectly.

1.2. Early learning of arithmetic as a source of difficulty

A growing body of work suggests that difficulties in understanding math equivalence may be largely attributable to children's early learning experiences in mathematics (Baroody & Ginsburg, 1983; Li, Ding, Capraro, & Capraro, 2008; McNeil, 2008; McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011). In the U.S., arithmetic problems are almost always presented in an "operations equals answer" format (e.g., $3+4=7$), which may fail to highlight the interchangeable nature of the two sides (McNeil et al., 2011; Seo & Ginsburg, 2003; but see Wynroth, 1975, as cited in Baroody & Ginsburg, 1983, for an atypical curriculum emphasizing relational meanings). As a result, many children come to interpret the equal sign *operationally*, as a signal to "give the answer," rather than *relationally*, as a signal that both sides share a common value (Baroody & Ginsburg, 1983; Behr et al., 1980; McNeil & Alibali, 2005a). Although operational interpretations of the equal sign are valid in some contexts (Seo & Ginsburg, 2003), they are inappropriate and often detrimental in algebraic contexts (Knuth et al., 2006; McNeil, Rittle-Johnson, Hattikudur, & Petersen, 2010). Consequently, most U.S. elementary school children not only fail to solve math equivalence problems correctly, but also fail to encode such problems' features correctly (McNeil & Alibali, 2004).

However, not all U.S. children exhibit such errors. What is it about the 10–25% of children who demonstrate understanding of math equivalence that enables them to extract appropriate patterns from their formal experiences with arithmetic? General competence or math ability alone cannot explain individual differences in understanding of math equivalence. Computational fluency, grade level, and age have not been consistently correlated with understanding of math equivalence in 7- to 11-year-olds. Some studies report no associations (Carpenter, Levi, & Farnsworth, 2000; McNeil &

Alibali, 2005b; McNeil et al., 2012; Rittle-Johnson & Alibali, 1999) and others a U-shaped (Baroody & Ginsburg, 1983; McNeil, 2007, 2008) or positive association (Rittle-Johnson, Matthews, Taylor, & McEldoon, 2011). Also, cross-cultural comparisons show understanding of math equivalence to relate more strongly to nationality than general competence (Capraro, Capraro, Ding, & Li, 2007).

Unfortunately, the subpopulation of children who succeed on math equivalence problems in the U.S. usually goes unstudied because their success excludes them from intervention studies. However, studying these individual differences may illuminate the mechanisms that lead to the development of a robust understanding of math equivalence and give educators tools for identifying children at risk for poor learning outcomes.

1.3. Knowledge organization as a source of individual differences

One factor that may affect the development of children's understanding of math equivalence is the way in which children's addition knowledge is organized. Gelman and Williams (1998) posited that when children organize their addition knowledge such that they link multiple instantiations of equivalent values (e.g., 2 + 4 = 6 and 1 + 1 + 1 + 1 + 1 + 1 = 6), it facilitates their ability to connect their addition knowledge to extant mathematical concepts. For example, multiple instantiations of six become linked and understood as interchangeable. Indeed, the structure of a person's domain knowledge is often as, if not more, influential than the content and quantity of that knowledge (Chase & Simon, 1973; Chi & Ceci, 1987).

Thus, differences in the ways children organize their addition knowledge may have profound effects on what they learn from their experiences with arithmetic. This idea is widely accepted by mathematics educators, who have long advocated helping children organize their knowledge of numbers into groupings based on equivalent values, particularly during the "New Math" movement of the 1960s and 1970s (Common Core State Standards Initiative, 2010; National Council of Teachers of Mathematics, 2000; School Mathematics Study Group, 1962; but see Kieran, 1980, for an alternative view).

Following this logic, we predicted that children who mentally organize addition knowledge in conceptually related groupings would have a better understanding of math equivalence than children who do not. Specifically, children who organize their addition knowledge based on equivalent values (e.g., 1 + 5, 3 + 3, 1 + 2 + 3, etc.) may come to understand the interchangeable nature of these facts more easily. A recent study provided initial support for this prediction (McNeil et al., 2012). Children aged 7–9 were randomly assigned to three conditions. In the experimental condition, children practiced addition problems grouped by equivalent values (i.e., 1 + 4, 2 + 3, 3 + 2, 4 + 1). In the "iterative" control condition, children practiced addition problems grouped in a standard iterative fashion (i.e., 2 + 1, 2 + 2, 2 + 3, 2 + 4). In the "no extra practice" control condition, children were given no practice with addition beyond what they normally received at home or at school. Children in the experimental condition showed a better understanding of math equivalence than children in the control conditions. McNeil et al. (2012) suggested that the experimental condition improved understanding of math equivalence by leading children to have a mental organization of addition knowledge that reflected equivalence. This conclusion was based, in part, on associationist models of knowledge in which the co-activation of math facts strengthens connections between them (McCloskey, Harley, & Sokol, 1991).

Although McNeil et al.'s (2012) study was consistent with the hypothesis that mental organization of addition knowledge by equivalent values supports understanding of math equivalence, it was lacking in two critical respects. First, the published paper did not include any data regarding children's mental organization of addition knowledge. Thus, it remains possible that the experimental intervention improved performance via some route other than changing children's mental organization of the number facts. For instance, it may have drawn attention to equivalent expressions in the external environment enough to boost understanding of math equivalence. Thus, no published data have shown that individual differences in mental organization of addition knowledge are associated with understanding of math equivalence.

Second, the McNeil et al. (2012) study is not able to establish that mental organization of addition knowledge is indeed associated with understanding of math equivalence in natural contexts. The experiment demonstrated that organizing addition facts by equivalent values in the *external* environment can improve children's understanding of math equivalence. However, it cannot speak to the question of whether children's *internal* organization of addition knowledge relates to children's understanding of math equivalence. The present study sought to address this question.

1.4. Present study

In two quasi-experiments, we used an individual differences approach to examine the relation between behavioral markers of the way children mentally organize their addition knowledge and their understanding of math equivalence. We expected children who show evidence of mentally organizing their addition knowledge into equivalent values to demonstrate a better understanding of math equivalence than children who do not show evidence of organizing their addition knowledge into such values. These quasi-experiments assessed children's mental organization of addition knowledge via two different methodologies.

In the first quasi-experiment, we investigated the prediction that children who use decomposition to solve addition problems would have a better understanding of math equivalence than children who did not, even when controlling for general proficiency with addition. Decomposition involves translating a problem into another known form to aid computation, typically breaking larger-valued numbers into more manageable values. Thus, this strategy integrates addition knowledge and solution procedures to allow solving strategies like "near-doubles" or "making-ten" (Baroody & Tiilikainen, 2003; Brownell, 1935; Cowan, 2003; Fayol & Thevenot, 2012; Rickard, 2005; Wilkins, Baroody, & Tiilikainen, 2001). For example, given 3 + 4, children may use a "near-doubles" decomposition strategy, extract the more easily solved 3 + 3 and simply add 1 to the result (Brownell & Chazal, 1935; Folsom, 1975; Rathmell, 1978; Siegler, 1987). Importantly, decomposition requires that knowledge be organized in a manner that allows one to access multiple equivalent representations of a given value. In order to decompose 4 into 3 + 1 (or 2 + 2, or 7 − 3), some mental connection between these equivalent representations must exist. It therefore seems reasonable that children who use a decomposition strategy are more likely to have knowledge organized around equivalent values than children who do not.

In the second quasi-experiment, we investigated the prediction that proficiency at generating sets of different equations equal to a target value relates to understanding of math equivalence. As with the decomposition strategy, in order to quickly and efficiently generate multiple equivalent representations (e.g., '8' ↔ '7 + 1', '4 + 4', '10 − 2', and so forth), some connection between representations must exist in memory (McCloskey et al., 1991). Children who can generate more equations that are equal to a requested value should therefore be more likely to have memory structures organized in conceptually related groupings than children who are less proficient at this task.

2. Quasi-experiment 1

2.1. Method

2.1.1. Participants

Participants were 101 children, ages 7–10; *M* age = 8.4, 56 girls. The group was 59% white, 27% African American or black, 10% Hispanic or Latino, 2% Asian, and 2% multiethnic. Most children (86%) were in grades 2–3 and between the ages of 7.5 and 9.5. (Others were between the ages of 7.0–7.5 and 9.6–10.5.) Age data were missing for 5 children, who were excluded from analyses that included age as a covariate. Conclusions were unchanged from analyses in which these children were included and age was not used as a covariate. Children were recruited from a diverse range of public and private elementary schools in a mid-sized city in the Midwestern United States. Approximately 50% received free or reduced price lunch. Each participant also served as a non-intervention control in one of three different studies on math equivalence (Chesney, McNeil, Petersen, & Dunwiddie, 2012; McNeil et al., 2012, 2011).

2.1.2. Procedure

Children participated individually in a 30-minute session. They completed a three-component measure of understanding of math equivalence followed by an addition strategy assessment.

2.1.3. Measures of understanding of math equivalence

All participants completed McNeil et al.'s (2011) three-component measure of understanding of math equivalence. The three components are: (a) equation solving, (b) equation encoding, and (c) defining the equal sign. According to McNeil and Alibali (2005b), these three components represent 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 participants.

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 did so, the experimenter asked, "Can you tell me how you got x?" (where x denotes the child's given answer).

Strategies were coded as correct or incorrect based on a system used in previous research (McNeil & Alibali, 2004; Perry, Church, & Goldin-Meadow, 1988). For most problems, correctness of the strategy could be inferred from the response itself (e.g., for the problem $3+5+6=3+\_$, a response of 11 indicated a correct strategy). Ambiguous responses were coded based on children's verbal explanation (e.g., a response of 9 to $3+5+6=3+\_$, was coded as incorrect if the child's verbal explanation was "I added 3 plus 5" but "correct" if the explanation was "I added 5 plus 6"). Agreement between coders on whether or not a given strategy was correct was 100%.

To assess *equation encoding*, children were videotaped as they reconstructed four math equivalence problems ($7+1=\_+6$, $3+5+4=\_+4$, $4+5=3+\_$, $2+3+6=2+\_$) after viewing each for 5 s (Chase & Simon, 1973; Siegler, 1976). The experimenter told children that they did not have to solve the problems, but rather that they needed only to write exactly what they saw after the experimenter displayed the problem. Encoding performance was coded as correct or incorrect based on a system used in previous research (McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999). Often, children erroneously converted problems to a traditional "operations equals answer" format (e.g., reconstructing $4+5=3+\_$ as "$4+5+3=\_$" or "$4+5=3$"). Agreement between coders on whether or not a reconstruction was correct was 100%.

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 on a piece of paper 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?" (Baroody & Ginsburg, 1983; Behr et al., 1980; Knuth et al., 2006). Children's responses were categorized according to a system used in previous research (McNeil & Alibali, 2005a, 2005b). 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", "the same as", "it means something is balanced"). Participants were coded as giving a relational definition of the equal sign if they gave a relational response to either question #2 or #3 above. Agreement between coders on whether or not a definition was relational was 100%.

2.1.4. Addition strategy assessment

Children were videotaped as they solved a set of 14 simple addition problems. Problems were presented one at a time on a computer screen. For each trial, a fixation point appeared in the center of the screen followed by the prompts "ready," "set," and "go," displayed for 1 s each. The two addends were then presented on the screen without an equal sign (e.g., $9+8$), and they remained on the screen until the child said his or her answer aloud. Reaction time (stimulus onset to verbal response) was recorded by the computer. After each trial, children were asked, "Can you tell me how you got that answer?" Children's strategies (following Geary, Bow-Thomas, Liu, & Siegler, 1996) were categorized as: (a) *finger counting*—child (when explaining how he or she arrived at the response) indicated using or was observed using his or her fingers to keep track of counts during problem solving, (b) *verbal counting*—child was observed counting aloud or indicated that he or she counted silently or softly during problem solving, (c) *retrieval*—child showed no evidence of finger counting or verbal counting and indicated that he or she "remembered" or "knew" the answer, or (d) *decomposition*—child indicated that he or she used a step-wise process involving a familiar addition fact (e.g., ten-based: $9+8=10+7=17$, tie-based: $7+8=7+7+1=14+1=15$) to derive the response. Children gave responses on over 99%

Table 1
Correlations of raw scores on all measures in Quasi-experiment 1.

	Decomposition use	Equation solving	Equation encoding	Defining the equal sign	Addition performance	Reaction time	Number of strategies
Solving	$r = .24^{*}$						
Encoding	$r = .30^{**}$	$r = .44^{***}$					
Defining	$r = .16$	$r = .27^{**}$	$r = .33^{***}$				
Addition	$r = .24^{*}$	$r = .24^{*}$	$r = .23^{*}$	$r = .05$			
RT	$r = -.27^{**}$	$r = -.21^{*}$	$r = -.27^{**}$	$r = -.09$	$r = -.60^{***}$		
Strategies	$r = .56^{***}$	$r = -.15$	$r = .21^{*}$	$r = .08$	$r = .18^{+}$	$r = -.33^{***}$	
Age ($N = 97$)	$r = -.04$	$r = .08$	$r = .28^{**}$	$r = .17$	$r = .14$	$r = -.15$	$r = .03$

$^{+}$ $p < .1$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.
$N = 101$ unless otherwise noted.

of the 14 trials, and all but 4 responses were codeable as uniquely belonging to one of these categories. Agreement between coders was 89% for categorizing children's strategies into one of the categories above; most of the disagreements occurred on strategies that were difficult to identify as finger counting versus verbal counting.

2.2. Results

2.2.1. Understanding of math equivalence

Consistent with previous research, children's performance was relatively poor. On average, children solved 0.58 ($SD = 1.28$) math equivalence problems correctly (of 4) and encoded 0.88 ($SD = 1.07$) math equivalence problems correctly (of 4). Only 9% of children provided a relational definition of the equal sign.

Because equation solving, equation encoding, and defining the equal sign are considered to be related components of children's understanding of math equivalence (McNeil & Alibali, 2005b), scores on the measures should be correlated with one another, and this was the case: equation solving and equation encoding, $r = .44$, $p < .001$; equation solving and defining the equal sign, $r = .27$, $p = .007$; equation encoding and defining the equal sign, $r = .33$, $p = .001$. Such correlations are consistent with previous research (McNeil & Alibali, 2004, 2005b; McNeil et al., 2011; Rittle-Johnson & Alibali, 1999) and provide some evidence of construct validity (Cronbach & Meehl, 1955). Correlations between raw scores on all measures are presented in Table 1.

2.2.2. Addition strategy assessment

Accuracy on the addition strategy assessment was near ceiling. On average, children solved 13.00 ($SD = 1.36$) simple addition problems correctly (of 14), and their response time (RT) was 5.64 s ($SD = 2.71$) per problem. Verbal counting was the most common solving strategy (35.5% of the trials), then finger counting (31.9%), retrieval (18.8%), and decomposition (13.8%).

2.2.3. Association between use of decomposition and understanding of math equivalence

Use of decomposition was relatively uncommon (13.8% of trials), and the data showed significant deviation from normality (Shapiro–Wilk, $p < .001$), with 58% of children never using decomposition. Thus, we analyzed the data by categorizing children into one of two groups: those who used decomposition at least once ($n = 42$), and those who did not ($n = 59$). Table 2 shows performance on each of the measures of understanding of math equivalence by these two decomposition groups. Because the pattern was similar for all three measures, we used a composite measure of math equivalence (McNeil et al., 2011; Rittle-Johnson & Alibali, 1999) for efficient presentation (Cohen, 1990). To take advantage of performance gradations while ensuring all three tasks were given equal weight, we summed

Table 2
Performance on each of the measures of understanding of math equivalence by decomposition use in Quasi-experiment 1.

Measure	Decomposition in repertoire	Decomposition not in repertoire
Equation solving (*M* out of 4 [*SD*])	0.95 (1.47)	0.32 (1.07)
Equation encoding (*M* out of 4 [*SD*])	1.26 (1.17)	0.61 (0.91)
Defining the equal sign (% who defined relationally)	14	5
Composite understanding (*M* sum of *z*-scores on the three measures)	0.83 (2.53)	−0.59 (1.83)

z-scores from each task to create the composite score. Scores on the composite measure ranged from −1.59 to 7.82 ($M = 0.00$, $SD = 2.25$).

We performed an analysis of covariance (ANCOVA) with decomposition use as the independent variable, with children's age and both accuracy and RT on the simple addition problems as covariates, and with score on the composite measure of understanding of math equivalence as the dependent measure. As predicted, there was a significant effect of decomposition use, $F(1, 92) = 6.38$, $p = .013$, $\eta^2 = .065$. Children who had decomposition in their strategy repertoires exhibited a better understanding of math equivalence ($M = 0.83$, $SD = 2.53$) than did children who did not ($M = -0.59$, $SD = 1.83$), *Cohen's d* = .66. The association seemed to be due to use of decomposition per se, as neither accuracy ($p = .60$) nor RT ($p = .36$) on the simple addition problems predicted understanding of math equivalence. Age correlated positively with understanding of math equivalence: $F(1, 92) = 5.23$, $p = .024$, $\eta^2 = .05$, $b = 0.72$, $SE = 0.32$, but the effect of decomposition use was maintained after controlling for age, *Cohen's d* = .63.

As scores were not normally distributed (Shapiro–Wilk, $p < .001$), we also performed a nonparametric analysis. We used binomial logistic regression to predict the log of the odds of scoring in the top half of the distribution (median split) on the composite measure of understanding of math equivalence, with decomposition use as the predictor and again controlling for children's age and for both accuracy and RT on the simple addition problems (Agresti, 1996). Results were consistent with the ANCOVA. The model estimated that the odds of scoring in the top half of the distribution on the measure were 2.87 times higher for children who had decomposition in their repertoires than for those who did not (30 of 42 [71%] versus 25 of 59 [42%]), $\beta = 1.05$, $SE = 0.47$, $Wald = 5.09$, $p = .02$. Neither accuracy ($p = .37$) nor RT ($p = .70$) on the simple addition problems was associated with understanding of math equivalence. Age was marginally associated with understanding of math equivalence $\beta = 0.61$, $SE = 0.34$, $Wald = 3.27$, $p = .070$.

2.2.4. Does use of other addition strategies predict understanding of math equivalence?

To determine if the observed effects were specific to the decomposition strategy, we repeated the parametric analyses reported above for each of the remaining strategies (verbal counting, finger counting, and retrieval). We found no evidence that the presence of these other strategies in a child's repertoire predicted understanding of math equivalence: verbal counting $F(1, 92) = 0.04$, $p = .84$; finger counting $F(1, 92) = 1.04$, $p = .31$; retrieval $F(1, 92) = 0.76$, $p = .38$. This suggests use of decomposition is unique in predicting children's understanding of math equivalence.

2.2.5. Does decomposition predict understanding when controlling for strategy variability?

Children who had decomposition in their strategy repertoires used a wider variety of strategies on the simple addition problems than did children who did not. An analysis of variance (ANOVA) with decomposition use as the independent variable, and number of strategies (of 4) in a child's repertoire as the dependent measure found a statistically significant effect of decomposition group, $F(1, 99) = 46.00$, $p < 001$, $\eta^2 = .32$. Children who had decomposition in their strategy repertoire used more strategies ($M = 3.17$, $SD = 0.82$) on the simple addition problems than did children who did not have decomposition in their repertoire ($M = 2.12$, $SD = 0.72$), *Cohen's d* = 1.37.

Because previous studies have documented relations between strategy variability and learning or performance, it is important to examine if the association between decomposition use and understanding of math equivalence holds when controlling for strategy variability. Thus, we repeated the original ANCOVA with decomposition group as the independent variable, with children's age and both accuracy and RT on the simple addition problems as covariates, and with score on the composite measure of understanding of math equivalence as the dependent measure, but included number of strategies in the children's repertoire as an additional covariate. Consistent with the previous ANCOVA, there was a statistically significant effect of decomposition use, $F(1, 91) = 4.89$, $p = .03$, $\eta^2 = .05$. None of the covariates were significant (accuracy $p = .61$, RT $p = .35$, and number of strategies in repertoire $p = .87$) with the exception of age, $F(1, 91) = 5.20$, $p = .02$, $\eta^2 = .05$, $b = 0.72$, $SE = 0.32$. These results suggest that children's decomposition use predicts their understanding of math equivalence and that this effect is not being driven by greater strategy variability.

2.3. Discussion

Children who used a decomposition strategy to solve addition problems scored higher on a composite measure of understanding of math equivalence than children who did not, even after controlling for other markers of addition proficiency and general competence (age, accuracy, reaction time, and strategy variability). As children who use the decomposition strategy show evidence of retrieving at least some addition knowledge in groups based on equivalent values (e.g., 3 + 4, 3 + 3 + 1), these results are consistent with our hypothesis that children who mentally organized their addition knowledge by common value would demonstrate better understanding of math equivalence.

However, there are at least two reasons to be cautious of this interpretation. First, while it is logical that use of decomposition is facilitated by mental organization of addition knowledge by equivalent values, the test of whether such mental organization is present is only indirect.

Decomposition use might also rely on an understanding that particular numbers and/or expressions can be substituted for one another. Although children's knowledge of which numbers and/or expressions are equal to one another says nothing per se about their understanding of the symbols used to denote this relation (i.e., the equal sign; Rittle-Johnson & Alibali, 1999), recent research indicates understanding of substitution is an important component of equal sign understanding (Jones, Inglis, Gilmore, & Dowens, 2012). Thus, it remains possible that decomposition use is simply another way to measure understanding of math equivalence.

Second, use of decomposition was relatively rare in this sample. Over half of the sample never used it at all, and nearly half of those who used it did so on less than a quarter of the trials. Only three participants used it on at least three quarters of the trials. We have suggested that the mere presence of decomposition in a child's repertoire provides evidence that the child has organized at least some of his/her addition knowledge by equivalent values. However, we have not assumed that a greater use of decomposition necessarily indicates greater levels of mental organization. Indeed, past work suggests that the availability of decomposition as a strategy does not automatically entail its use, particularly when other, potentially less effortful strategies— such as direct retrieval—are also available (see Fayol & Thevenot, 2012 for a discussion of strategy variability in supposed experts). To elaborate, decomposition requires organized knowledge of related arithmetic facts. Thus, knowledge of more arithmetic facts would increase the proportion of problems on which one could use decomposition, provided that that knowledge is organized by equivalent values. However, knowledge of more arithmetic facts would *also* increase the availability of the retrieval strategy, which does not entail any particular mental organization. Consistent with this view, our results showed a strong relation between use of decomposition and understanding of math equivalence, although they provided no indication that frequency of decomposition use mattered. However, the data were not sufficiently well-distributed to test if frequency actually mattered.

To address concerns about the appropriateness of decomposition as a metric of mental organization, we conducted a second quasi-experiment in which we used a different method to assess children's mental organization of addition knowledge—one that focused on recall more directly, rather than on decomposition.

3. Quasi-experiment 2

In Quasi-experiment 2, we sought to replicate the finding that the mental organization of addition knowledge is associated with understanding of math equivalence. However, we assessed mental organization by children's ability to generate equations equal to a target value. The ability to generate equivalent equations should benefit from mental organization that links mathematical statements with equivalent values. We predicted that proficiency at generating equations with equal values would be associated with higher scores on our assessment of understanding of math equivalence, even after controlling for general addition proficiency.

3.1. Method

3.1.1. Participants

The quasi-experiment was conducted in second grade classrooms at two public and two private schools in a mid-sized city in the Midwestern US and one private school in a large city in the Eastern US. Students in these classes were part of a larger study designed to examine the effects of different types of arithmetic practice on children's understanding of math equivalence. Children solved problems in arithmetic practice workbooks 15 min per day, twice a week, for 12 weeks, during their regular mathematics period. The specific problems of interest for the present study were contained in workbooks that were randomly assigned to approximately half of the children in each classroom (the remaining children completed different workbooks not relevant to testing our current hypothesis). Ninety-four children (50 girls) completed all the problems of interest. Five additional children were excluded because they were absent on one or more days of testing. Based on the composition of the participating schools, the racial/ethnic makeup of the sample was approximately 48% white, 23% African American or black, 19% Hispanic or Latino, 2% Asian, and 8% other. Approximately 50% received free or reduced price lunch.

3.1.2. Measure of proficiency at generating a set of equations equal to a target value

Children completed a series of problems designed to test their ability to generate a set of equations equal to a target value. Each problem displayed a cartoon character ("Mary" or "Juan") along with a statement indicating that she or he "likes" a particular target number. An equation board containing 2–4 blank slots was presented below the character. The children's goal was to fill the empty equation board with equations that the character would "like." (See Fig. 1.)

Children completed 20 of these empty equation board problems across two sets of worksheets presented 1–2 weeks apart. There were a total of 72 blank equation board slots that children could fill with equations. Prior to completing each set, children received instructions from their classroom teacher and then worked through a scaffolded example (see Fig. 1). In addition, for each set of worksheets the first four equation boards that children completed on their own contained 1 or 2 example responses to ensure that children understood the task. The goal was to fill in as many blank equation board slots as possible in 15 min (~30 s per blank equation board slot).

The "equation board" worksheets were embedded in the larger arithmetic practice workbook that children used in their regular mathematics classrooms as part of the larger study. This workbook provided practice with single-digit addition facts in a variety of formats. Teachers were encouraged to treat these sessions as any other math practice activity they would assign and to answer children's questions accordingly.

Children were given one point for every correct equation or expression they wrote, as long as it did not exactly replicate an equation or expression written in an earlier slot on the same board. For example, if a child correctly filled in the first slot in a board with "3 + 4 = 7," subsequent entries of "4 + 3 = 7," "7 = 5 + 2," and "8 – 1" would each be scored as correct, whereas a duplicate entry of "3 + 4 = 7" or "3 + 4" would not. We established inter-rater reliability by having a second coder code the responses. Agreement between two coders on whether or not a given equation or expression was correct was 99%. The coders resolved all disagreements through discussion.

3.1.3. Measures of understanding of math equivalence

As in Quasi-experiment 1, children's understanding of math equivalence was evaluated using a three-component measure assessing equation solving, equation encoding, and defining the equal sign. Minor adjustments were made to the procedure to allow classroom collection of the data,

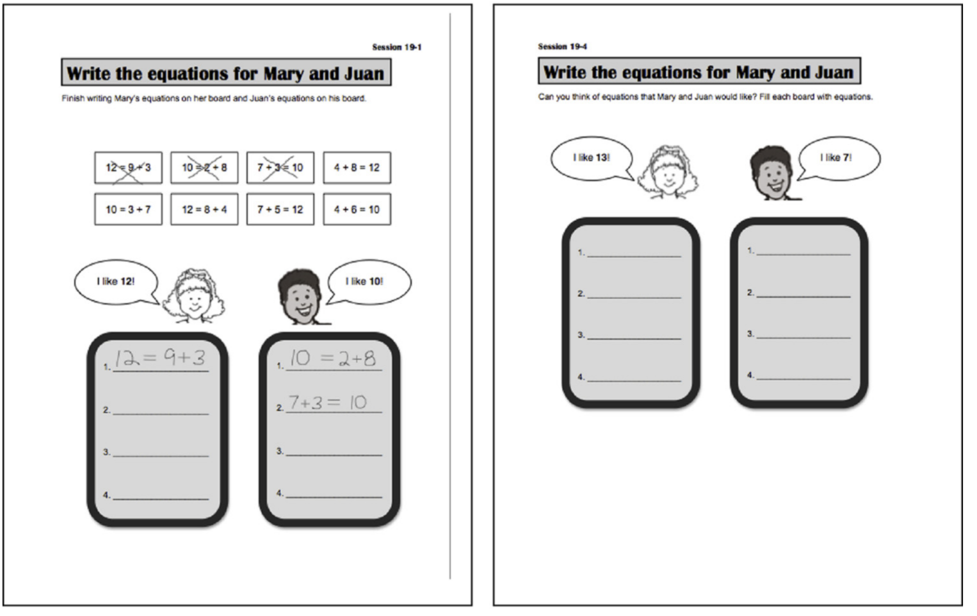

Fig. 1. Sample pages from the workbook activities used in Quasi-experiment 2. The left panel shows the partially completed scaffolded example children completed before filling out equation boards on their own. The right panel shows a worksheet with empty equation boards.

noted below. A second coder coded responses of 20% of children to establish inter-rater reliability. Agreement between coders is reported below for each component.

To assess *equation solving*, children completed a paper-and-pencil test consisting of eight math equivalence problems (5 + 4 = _ + 4; 8 + 2 = _ + 7; 7 + 2 + 4 = _ + 4; 7 + 4 + 6 = _ +3; 2 + 6 = 2 + _; 2 + 7 = 6 + _; 3 + 5 + 6 = 3 + _; 6 + 2 + 8 = 4 + _). Children took the test in their regular classroom setting. Children's classroom teachers handed out the test and read the instructions aloud. Teachers assured children that it was okay if they had never seen problems like these before. They told children to try their best to solve each problem and to give their "best guess" for any problems that seemed too difficult. As in prior work involving paper-and-pencil assessment of strategy use on these types of equations (McNeil, 2007), responses were coded as reflecting a correct strategy as long as they were within $\pm$1 of correct. 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) using paper and pencil after viewing each for 5 s (Chase & Simon, 1973; Siegler, 1976). Teachers read the following instructions aloud: "I will show you some math problems one at a time, but this time you don't have to solve the problems. Instead, you just have to remember what you see and write it on your paper. So, here's what we'll do. I'll show you a problem for 5 s. *After* I hide the problem, I want you to write exactly what you saw. Remember, you don't have to solve the problem; you just need to write exactly what you see." Teachers projected each equation using an overhead projector and left it visible for 5 s. After the teachers hid each problem, they said: "OK, write exactly what you saw." Children's encoding performance was coded as correct or incorrect via the same method used in Quasi-experiment 1. Agreement between coders on whether or not a given reconstruction was correct was 99%.

To assess *defining the equal sign*, children responded to a set of questions about the equal sign. An arrow pointed to an equal sign presented alone, and the text read: (1) "What is the name of this math symbol?" (2) "What does this math symbol mean?" and (3) "Can it mean anything else?" (Baroody & Ginsburg, 1983; Behr et al., 1980; Knuth et al., 2006). Teachers read each question aloud to the

Table 3
Performance on each of the measures of understanding of math equivalence for children with high versus low proficiency at generating equations equal to a target value in Quasi-experiment 2.

Measure	High proficiency	Low proficiency
Equation solving (*M* out of 8 [*SD*])	2.33 (3.13)	2.02 (2.82)
Equation encoding (*M* out of 4 [*SD*])	1.88 (1.21)	1.00 (1.12)
Defining the equal sign (% who defined relationally)	23	11
Composite understanding (*M* sum of *z*-scores on the three measures)	0.55 (2.02)	−0.58 (1.85)

Note. Proficiency at generating equations equal to a target value was median split for illustrative purposes in this table.

class, and children wrote their responses immediately after the question was read. Responses were categorized via the same method used in Quasi-experiment 1. Agreement between coders on whether or not a given definition was relational was 100%.

3.1.4. Measure of general addition proficiency

Children completed Geary et al.'s (1996) paper-and-pencil addition test. It contained all pair-wise combinations of the numbers 1–9, for a total of 81 addition problems, presented in a random order. The goal was to solve as many problems as possible in 1 min.

3.1.5. Procedure

The measures for this quasi-experiment were administered over a 3–5 week period in late fall. Children's teachers administered all measures in the regular mathematics classroom setting. Teachers were given some flexibility regarding the exact dates and times of administration to suit their classroom schedules; however, all children completed the measures according to the same general timetable. First, they completed the initial set of worksheets assessing proficiency at generating equations equal to target values. Then, approximately 1–2 weeks later, they completed the second set of equation generation worksheets. Finally, approximately 1–2 weeks after completing the second set of worksheets, children completed both the three-component composite measure of understanding of math equivalence and the measure of general addition proficiency.

3.2. Results

3.2.1. Proficiency at generating equations equal to a target value

Performance on the equation generation worksheets was highly varied. On average, children generated 53.02 (*SD* = 22.39, min = 0, max = 72) different correct equations to fill the equation boards across both sets of worksheets (of 72 possible).

3.2.2. Understanding of math equivalence

Children's understanding of math equivalence was relatively poor. On average, children solved 2.18 (*SD* = 2.97) math equivalence problems correctly (of 8) and encoded 1.45 (*SD* = 1.24) math equivalence problems correctly (of 4). Only 17% of children provided a relational definition of the equal sign.

3.2.3. General addition proficiency

Performance on the paper-and-pencil addition test was highly varied. On average, children solved 14.49 (*SD* = 6.49, min = 2, max = 32) problems correctly in 1 min (of 81 possible).

3.2.4. Association between proficiency at generating equations equal to a target value and understanding of math equivalence

Table 3 presents performance on each component of the math equivalence measure for children who were more and less proficient at generating a set of equations equal to a target value. The effect was similar across all three components, suggesting it was appropriate to use a composite measure of understanding of math equivalence. Scores on the composite measure ranged from −2.35 to 5.07 (*M* = 0.00, *SD* = 2.01). Correlations between raw scores on all measures are presented in Table 4.

Table 4
Correlations of raw scores on all measures in Quasi-experiment 2.

	Generating equations	Equation solving	Equation encoding	Defining the equal sign
Solving	$r = .13$			
Encoding	$r = .33^{**}$	$r = .22^{*}$		
Defining	$r = .19^{+}$	$r = .16$	$r = .13$	
Addition proficiency	$r = .37^{***}$	$r = .06$	$r = .23^{*}$	$r = .14$

$^{+}$ $p < .1$.
* $p < .05$.
** $p < .01$.
*** $p < .001$
$N = 94$.

We used hierarchical multiple regression to predict children's scores on the composite measure. In the first step, we entered scores on the paper-and-pencil addition test to control for general addition proficiency. There was an association between children's addition proficiency and their understanding of math equivalence (see Table 5). The more proficient children were with addition, the better their understanding of math equivalence, $b = 0.066$, $t(92) = 2.10$, $p = .04$. This effect was in the small-to-medium range, with general addition proficiency accounting for less than 5% of the variance in understanding of math equivalence when no other variables were included in the model (corresponding *Cohen's d* $\approx .45$).

In the second step, we entered the predictor variable of interest into the model: proficiency at generating equations equal to a target value. This predictor accounted for a significant portion of the variance in children's understanding of math equivalence, even when controlling for general addition proficiency (see Table 5). As hypothesized, the more proficient children were at generating such equations, the better their understanding of math equivalence, $b = 0.025$, $t(91) = 2.65$, $p = .01$. The effect was in the medium range, with proficiency at generating equations equal to a target value uniquely accounting for over 7% of the variance in understanding of math equivalence (corresponding *Cohen's d* $\approx .56$). Importantly, when we reversed the steps of the hierarchical regression and entered proficiency at generating equations first, equation generation accounted for 10% of the variance ($b = 0.03$, $t(92) = 3.26$, $p = .002$, corresponding *Cohen's d* $\approx .67$) and the association between general addition proficiency and understanding of math equivalence reported in the previous paragraph was no longer significant (R^2 change = .01, $p = .30$). Thus, consistent with Quasi-experiment 1, general addition proficiency was not a robust predictor of understanding of math equivalence, whereas a measure assessing the organization of addition knowledge was.

Given that scores on the composite measure of understanding of math equivalence were not normally distributed (Shapiro–Wilk, $p < .001$), we also performed a nonparametric analysis. We used binomial logistic regression to predict the log of the odds of scoring in the top half of the distribution (median split) on the composite measure of understanding of math equivalence, with proficiency at generating equations as the primary predictor and again controlling for general addition proficiency (Agresti, 1996). Results were consistent with the original hierarchical multiple regression. As predicted, children's likelihood of scoring in the top half of the distribution on the composite measure of understanding of math equivalence increased as their proficiency at generating equations increased,

Table 5
Predicting scores on the measure of understanding of math equivalence by the number of equations equal to a target value generated via a stepwise linear regression in Quasi-experiment 2.

Predictor	b	pr	t	p	R^2 change
Step 1				.04	.046
Score on paper-and-pencil addition test	0.066	.21	2.10	.04	
Step 2				< .01	.068
Score on paper-and-pencil addition test	0.034	.11	1.04	.30	
Number of equations equal to a target value generated	0.025	.27	2.65	.01	

$\beta = 0.03$, $z = 2.45$, $Wald = 6.09$, $p = .01$. The model indicated that the odds of scoring in the top half of the distribution increase by 81% for every standard deviation increase in equation generation.

3.3. Discussion

Children who were more proficient at generating equations equal to a target value also performed better on the math equivalence assessment, even after controlling for general addition proficiency. We suspect that children who are better able to generate such equations tend to mentally organize addition knowledge by equivalent values. This organization, in turn, may help children better understand the interchangeable nature of these equivalent values, and understanding this interchangeability is key to understanding math equivalence. Just as practice with a given addition fact strengthens associations between a pair of addends and their value (e.g., 4 + 5 activates 9), mentally organizing addition knowledge based on equivalent values may strengthen associations between multiple equivalent combinations. For example, 4 + 5 may come to activate not only the number 9 but also multiple combinations that equal 9 (e.g., 5 + 4, 3 + 6, 2 + 7). Then, when faced with a math equivalence problem such as 4 + 5 = _ + 6, children who organize their knowledge in this way may see 4 + 5 and immediately activate 3 + 6.

4. General discussion

The present study provides evidence that the organization of children's addition knowledge may be an important source of individual differences in children's understanding of math equivalence. In Quasi-experiment 1, children who used decomposition when solving addition problems exhibited a better understanding of math equivalence than children who did not use decomposition, even after controlling for age, accuracy, reaction time, and strategy variability. In Quasi-experiment 2, children who were able to generate more different equivalent equations demonstrated a better understanding of math equivalence than children who generated fewer equivalent equations, even after controlling for addition proficiency.

These results complement and extend the causal evidence established by McNeil et al. (2012) in at least two ways. First, they show that individual differences in children's addition knowledge are associated with understanding of math equivalence in symbolic form. Second, they show this association in children whose only experience with mathematics has been that of a typical U.S. elementary school student, not just in children who have completed interventions designed to influence their mental organization of addition knowledge. Moreover, the proxies for mental organization of addition knowledge predict understanding of math equivalence even when controlling for age and addition performance, RT, and strategy variability (Quasi-experiment 1), and general addition proficiency (Quasi-experiment 2).

4.1. Potential mechanisms

While our evidence indicates that the mental organization of addition knowledge predicts children's understanding of math equivalence, the mechanisms underlying this effect are less clear. Gelman and Williams' (1998) explanation of how the structure of learning environments can structure the mind offers a plausible mechanism. They argued that with exposure to more structurally redundant examples of domain-relevant exemplars, it becomes increasingly likely that a learner's mental structures will become more compatible with the underlying structure of the domain. Unpacking this theory, we offer speculative accounts of how the organization of addition knowledge might contribute to understanding of math equivalence.

It may be that conceptually grouped organization of addition knowledge leads to the co-activation of multiple interchangeable number combinations when a given stimulus is presented. In this way, co-activation could help populate a child's strategy space when faced with difficult or unfamiliar problems. For instance, when presented with 9 + 3 = 6 + _, a variety of equivalent forms might be activated that aid success with math equivalence problems (e.g., 6 + 3 + 3 = 6 + _). This co-activation has several possible consequences.

One is increased strategy choices. Siegler (1996) has argued cogently that the population of strategy space is a critical component of cognitive change. As discussed above, organizing math facts by equivalent values facilitates decomposition strategies. Thus, conceptual organization around equivalent values may help drive development of domain knowledge—i.e., understanding of math equivalence—because it increases children's strategy options.

In addition, co-activation of equivalent number combinations may counteract the operational default problem-solving strategy to "add up all the numbers" (McNeil & Alibali, 2005b). Both decomposition and equivalent equation generation work in the opposite direction; instead of adding all the numbers up, they are broken apart and rearranged. This availability of strategies for substituting multiple terms in place of any given term may work against an "add all" strategy, both by creating competition within the strategy space and by challenging its status as a default.

It may be that the organization of addition knowledge around equivalent groups helps children to directly extract conceptual knowledge about both number and math equivalence. According to Gelman and Williams (1998), structural mapping is facilitated when learners are exposed to multiple examples that vary in their surface details as long as their shared structure maps to the to-be-learned conceptual domain. Organizing addition knowledge by values fits this description, as each combination varies in surface appearance but remains equivalent. This interpretation aligns with Chi and Ceci's (1987) argument that quantity of knowledge is often not as important as knowledge structure. That is, content knowledge—and its organization—are not mere adjuncts of cognition, but are constituents of it. As such, representations of arithmetic knowledge that allow for the integration of factual and conceptual knowledge by supporting multiple concept-based interconnections may be more conducive to conceptual development than those based on isolated connections between ordered addend pairs and their sums (Brownell, 1935; Wilkins et al., 2001).

There are reasons to be cautious when interpreting the present results. First, it is possible that children initially gain an understanding of math equivalence and only afterwards come to mentally organize addition knowledge by equivalent values (Canobi, Reeve, & Pattison, 1998). Alternatively, mental organization and knowledge of math equivalence may reinforce each other iteratively (Rittle-Johnson, Siegler, & Alibali, 2001). Further studies are needed to determine if mental organization of addition knowledge by equivalent values precedes, follows, or develops in concert with understanding of math equivalence.

Second, we cannot currently measure mental organization of addition knowledge directly; we can only infer this structure indirectly. While the use of different behavioral indicators of mental organization lends strength to our conclusions (Lykken, 1968), we cannot rule out the possibility that the associations we found are due to their common relation to some unmeasured third factor. This concern is mitigated somewhat by McNeil et al.'s (2012) experiment, in which a random subset of children was assigned to an intervention with organized arithmetic practice designed to strengthen their mental connections between equivalent math facts. Random assignment should have yielded groups that were roughly equivalent on unmeasured third factors. Thus, it is valid to presume that the posttest differences in children's understanding of math equivalence resulted from the intervention (i.e., from the organized arithmetic practice) rather than from inherent differences between the children. Given the converging lines of evidence from McNeil et al.'s (2012) experiment and both of the current quasi-experiments, the most parsimonious interpretation of these results is a causal connection between organization of addition knowledge based on equivalent values and understanding of math equivalence. However, we acknowledge that the available evidence is not definitive.

Third, it remains an open question why some children but not others would come to organize addition knowledge by equivalent values. One possibility is chance. Random variation in co-exposure and momentary attention may leave some children with stronger interconnections between representations with equivalent values than others. A second possibility is that these children had different mathematical experiences than their peers (e.g., math workbooks, computer games, or flashcards used at home) that enabled them to construct a better understanding of math equivalence. A third possibility is that there is some intrinsic difference in the general cognitive ability of these children that facilitated the development of memory structures that interconnect equivalent addition facts (Geary, Hoard, & Nugent, 2012). Further study is needed to determine which possibility, if any, is the case.

In conclusion, our results suggest children who organize their addition knowledge into groupings based on equivalent values construct a better understanding of math equivalence than children who do not. These findings run counter to Kieran's (1980) conjecture that emphasizing the equivalence of arithmetic facts in classrooms would lead to poor learning of those facts because children have poor understanding of equivalence. Instead, our findings bolster the recommendations of math educators who have endorsed helping children organize their knowledge of numbers and math facts into groupings based on equivalent values (Common Core State Standards Initiative, 2010; School Mathematics Study Group, 1962). Emphasizing these interconnections may improve understanding of math equivalence in symbolic form without detriment to arithmetic fact learning. Beyond these practical implications, the present results also provide an illustration of how examining individual differences can help illuminate the mechanisms involved in mathematical thinking.

References

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

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. http://dx.doi.org/10.1086/461356

Baroody, A. J., & Tiilikainen, S. H. (2003). Two perspectives on addition development. In A. J. Baroody, & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 75–125). Mahwah, NJ: Erlbaum.

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

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing*, *11*, 239–273. http://dx.doi.org/10.1023/A:1008050403932

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *41*, 189–201. http://dx.doi.org/10.1037/0012-1649.41.6.189

Brownell, W. A. (1935). Psychological considerations in the learning and the teaching of arithmetic. In D. W. Reeve (Ed.), *The teaching of arithmetic: Tenth yearbook of the national council of teachers of mathematics* (pp. 1–31). New York: Columbia University Press.

Brownell, W. A., & Chazal, C. B. (1935). The effects of premature drill in third-grade arithmetic. *Journal of Educational Research*, *29*, 17–28.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology*, *34*, 882–891. http://dx.doi.org/10.1037/0012-1649.34.5.882

Capraro, R. M., Capraro, M. M., Ding, M., & Li, X. (2007). Thirty years of research: Interpretations of the equal sign in China and the USA. *Psychological Reports*, *101*, 784–786. http://dx.doi.org/10.2466/pr0.101.3.784-786

Carpenter, T. P., Levi, L., & Farnsworth, V. (2000). *Building a foundation for learning algebra in the elementary grades*. Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics Education.

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

Chesney, D. L., McNeil, N. M., Petersen, L. A., & Dunwiddie, A. E. (2012). Arithmetic practice that includes relational words promotes conceptual understanding and computational fluency. In *Poster presented at the annual convention for the Association for Psychological Sciences* Chicago, IL.

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. http://dx.doi.org/10.1016/S0065-2407(08)60401-2

Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, *45*, 1304–1312. http://dx.doi.org/10.1037/0003-066X.45.12.1304

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices and Council of Chief State School Officers.

Cowan, R. (2003). Does it all add up? Changes in children's knowledge of addition combinations, strategies, and principles. In A. J. Baroody, & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 35–74). Mahwah, NJ: Erlbaum.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *30*, 671–684. http://dx.doi.org/10.1037/h0043943

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302. http://dx.doi.org/10.1037/h0040957

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, *6*, 232–236.

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, *123*, 392–403. http://dx.doi.org/10.1016/j.cognition.2012.02.008

Folsom, M. (1975). Operations on whole numbers. In J. N. Payne (Ed.), *Mathematics learning in early childhood (1975 yearbook of the National Council of Teachers of Mathematics)* (pp. 161–190). Reston, VA: NCTM.

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, schooling. *Child Development*, *67*, 2022–2044. http://dx.doi.org/10.2307/1131607

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology*, *113*, 49–65. http://dx.doi.org/10.1016/j.jecp.2012.03.003

Gelman, R., & Williams, E. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In D. Kuhn, W. Damon, R. Siegler, & R. Lerner (Eds.), *Handbook of child psychology. Vol. 2: Cognition, perception, and language* (5th ed., Vol. 20, pp. 575–630). New York: John Wiley and Sons.

Hughes, C., Jaffee, S. R., Happe, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in theory of mind: From nature to nurture? *Child Development*, *76*, 356–370. http://dx.doi.org/10.1111/j.1467-8624.2005.00850_a.x

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.

Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: Two components of a relational conception of the equals sign. *Journal of Experimental Child Psychology*, *113*, 166–176. http://dx.doi.org/10.1016/j.jecp.2012.05.003

Kieran, C. (1980). The interpretation of the equal sign: Symbol for an equivalence relation vs. an operator symbol. In R. Karplus (Ed.), *Proceedings of the fourth conference of the international group for the psychology of mathematics education* (pp. 163–169). Berkeley, CA: University of California.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, *12*, 317–326. http://dx.doi.org/10.1007/BF00311062

Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390–419). New York: Macmillan.

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.

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. http://dx.doi.org/10.1080/07370000801980845

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*, 151–159. http://dx.doi.org/10.1037/h0026141

McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 377–397. http://dx.doi.org/10.1037/0278-7393.17.3.377

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

McNeil, N. M. (2007). U-shaped development in math: 7-Year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, *43*, 687–695. http://dx.doi.org/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 mathematical equivalence. *Child Development*, *79*, 1524–1537. http://dx.doi.org/10.1111/j.1467-8624.2008.01203.x

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. http://dx.doi.org/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. http://dx.doi.org/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. http://dx.doi.org/10.1111/j.1467-8624.2005.00884.x

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., et al. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. *Journal of Educational Psychology*, *104*, 1109–1121. http://dx.doi.org/10.1037/a0028997

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 mathematical equivalence. *Child Development*, *82*, 1620–1633. http://dx.doi.org/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*, 437–457. http://dx.doi.org/10.1080/15248372.2010.516421

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

Nelson, K. (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology*, *17*, 170–187. http://dx.doi.org/10.1037/0012-1649.17.2.170

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development*, *6*, 449–468. http://dx.doi.org/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. http://dx.doi.org/10.1016/0885-2014(88)90021-4

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. http://dx.doi.org/10.1111/j.1467-8624.2007.01131.x

Rathmell, E. C. (1978). Using thinking strategies to teach the basic facts. In M. Suydam (Ed.), *Developing computational skills (1975 yearbook of the National Council of Teachers of Mathematics)* (pp. 13–38). Reston, VA: NCTM.

Rickard, T. C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 250–257. http://dx.doi.org/10.1037/0278-7393.31.2.250

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. http://dx.doi.org/10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, *103*, 85–104. http://dx.doi.org/10.1037/a0021334

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*, 346–362. http://dx.doi.org/10.1037/0022-0663.93.2.346

School Mathematics Study Group. (1962). *Mathematics for the elementary school: Grade 6*. New Haven, CT: Yale University Press.

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.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, R. K., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*, 926–933. http://dx.doi.org/10.1016/j.biopsych.2003.12.019

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

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264. http://dx.doi.org/10.1037/0096-3445.116.3.250

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

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *Psychological Review*, *66*, 297–333. http://dx.doi.org/10.1037/h0076759

Wilkins, J. L. M., Baroody, A. J., & Tiilikainen, S. (2001). Kindergartners' understanding of additive commutativity within the context of word problems. *Journal of Experimental Child Psychology*, *79*, 23–36. http://dx.doi.org/10.1006/jecp.2000.2580