OPTIMIZING PROBLEM FORMAT TO FACILITATE CHILDREN’S UNDERSTANDING OF MATH EQUIVALENCE

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Caroline Byrd Hornburg

Nicole M. McNeil, Director

Graduate Program in Psychology
Notre Dame, Indiana

April 2017
Mathematical equivalence is a foundational concept. Unfortunately, most children (ages 7-11) struggle to understand it and have difficulties solving *math equivalence problems*, which have operations on both sides of the equal sign (e.g., \(3 + 4 = \_ + 5\)). One prevailing explanation for children’s poor performance is that children encode, or internally represent, the problems inaccurately. Previous research has found that children’s encoding of math equivalence problems depends on format, with accuracy encoding right-blank problems (e.g., \(3 + 4 = 5 + \_\)) being much lower than accuracy encoding left-blank problems (e.g., \(3 + 4 = \_ + 5\)). Following the prevailing account, if poor encoding causes solving difficulties, then right-blank problems should be more difficult to solve correctly than left-blank problems. However, a second, equally plausible prediction is that the role of encoding in solving differs based on problem format. This project examined the role of encoding in (1) children’s performance solving math equivalence problems, using an integrative data analysis, and (2) children’s learning to solve math equivalence problems correctly, using a randomized experiment. Contrary to
the prevailing view that encoding performance predicts solving performance, accuracy on right-blank problems was significantly higher than accuracy on left-blank problems. Additionally, overall encoding accuracy was significantly correlated with performance on right-blank problems only. Results indicate that the role of encoding in children’s solving of math equivalence problems depends on problem format, and specifically that the interpretation of what is encoded may matter more than accuracy encoding problem features. Furthermore, results from the randomized experiment suggest it may be more efficient to teach math equivalence using right-blank versus left-blank problems.
CONTENTS

Figures.............................................................................................................. iv

Tables .............................................................................................................. v

Acknowledgments ......................................................................................... vi

Chapter 1: Introduction .................................................................................. 1
  1.1 Links Between Problem Encoding and Problem Solving ...................... 5
  1.2 The Change-Resistance Account ............................................................ 7
  1.3 Alternative Accounts of Children’s Performance .................................. 9
     1.3.1 Run-on Schema ........................................................................... 10
     1.3.2 Difficult Unknown ....................................................................... 11
  1.4 The Role of Encoding in Children’s Performance ................................ 12
  1.5 The Role of Encoding in Children’s Learning ...................................... 13

Chapter 2: Study 1 – Examining the Role of Encoding in Children’s Performance .... 17
  2.1 Method .................................................................................................... 17
     2.1.1 Design ......................................................................................... 17
     2.1.2 Participants .................................................................................. 18
     2.1.3 Measures .................................................................................... 19
        2.1.3.1 Problem Solving .................................................................. 19
        2.1.3.2 Problem Encoding ................................................................. 21
  2.2 Data Analytic Plan .................................................................................. 21
     2.2.1 Unconditional Models .................................................................. 22
     2.2.2 Conditional Models ..................................................................... 23
  2.3 Results .................................................................................................... 24
     2.3.1 Overall Performance .................................................................... 24
     2.3.2 Unconditional Models of Performance by Problem Format .......... 24
     2.3.3 Conditional Models of Performance by Problem Format ............ 27
     2.3.4 Relations Between Solving and Encoding .................................... 28
     2.3.5 Incorrect Strategy Analyses ......................................................... 29
     2.3.6 Arithmetic Error Analyses ............................................................ 30
  2.4 Discussion ............................................................................................... 30
     2.4.1 Theoretical Implications ............................................................... 31
     2.4.2 Practical Implications .................................................................. 33
     2.4.3 Limitations and Future Directions .............................................. 34
  2.5 Figures .................................................................................................... 37
  2.6 Tables ...................................................................................................... 40
Chapter 3: Study 2 – Examining the Role of Encoding in Children’s Learning

3.1 Method...........................................................................................................48
  3.1.1 Participants.........................................................................................48
  3.1.2 Measures .............................................................................................49
    3.1.2.1 Pretest .........................................................................................49
    3.1.2.2 Task Difficulty Rating .................................................................51
    3.1.2.3 Posttest, Opposite Format, and Other Format Problems..52
  3.1.3 Procedure ............................................................................................54
    3.1.3.1 Experimental Conditions ...........................................................54
  3.2 Results......................................................................................................56
    3.2.1 Pretest Performance ........................................................................56
    3.2.2 Accuracy During and After the Intervention by Test Phase and
       Condition .........................................................................................57
    3.2.3 Does Pretest Encoding Accuracy Moderate the Effect of the
       Interventions? ..................................................................................60
    3.2.4 Levels of Scaffolding Needed to Learn ..........................................61
    3.2.5 Do Children Maintain and Transfer What They Learned in the
       Intervention? ....................................................................................64
    3.2.6 Task Difficulty Rating by Problem Format ...................................66
    3.2.7 Post-hoc Analyses .............................................................................67
  3.3 Discussion ..............................................................................................69
    3.3.1 Theoretical Implications ..................................................................69
    3.3.2 Practical Implications ......................................................................73
    3.3.3 Limitations and Future Directions ..................................................74
  3.4 Figures ....................................................................................................78
  3.5 Tables .....................................................................................................86

Chapter 4: General Discussion......................................................................90
  4.1 Conclusion ..............................................................................................94

Appendix A: Randomized Experiment: Session Script..................................95

References......................................................................................................109
FIGURES

Figure 2.1: Accuracy solving left-blank (top graph) and right-blank problems (bottom graph) for the full sample ($N = 1,414$).................................37

Figure 2.2: Accuracy solving left-blank (top graph) and right-blank problems (bottom graph) for the subsample ($n = 196$)...................................................38

Figure 2.3: Overall encoding accuracy for the full sample (top graph; $N = 981$) and the subsample (bottom graph; $n = 133$)...................................................39

Figure 3.1: Task difficulty rating for the experimental conditions (top panel) and control condition (bottom panel)...............................78

Figure 3.2: Levels of scaffolding................................................................79

Figure 3.3: Adjusted problem-solving accuracy across test phases by condition........80

Figure 3.4: Unadjusted problem-solving accuracy by experimental condition (left-blank, $n = 36$; right-blank, $n = 34$) for intervention problems (top row) and posttest problems (bottom row)........................................................81

Figure 3.5: Unadjusted problem-solving accuracy by experimental condition (left-blank, $n = 36$; right-blank, $n = 34$) for opposite format problems (top row) and other format problems (bottom row)........................................................82

Figure 3.6: Adjusted posttest encoding accuracy for each problem format by condition. 83

Figure 3.7: Unadjusted posttest encoding accuracy by condition (left-blank, $n = 36$; right-blank, $n = 34$; control, $n = 35$) for intervention (left column) and opposite problems (right column)........................................................84

Figure 3.8: Percent of children learning at each level of scaffolding by condition. 85
Table 2.1 Integrative Data Analysis: Study Characteristics ........................................40
Table 2.2 Integrative Data Analysis: Problems Given in Each Study .................41
Table 2.3 Example Problem-Solving Strategies and Explanations ....................43
Table 2.4 Integrative Data Analysis: Problem Solving Performance by Study ..........44
Table 2.5 Integrative Data Analysis: Problem Encoding Performance by Study ..........45
Table 2.6 Integrative Data Analysis: Predicting Cohen’s d .......................................46
Table 2.7 Integrative Data Analysis: Predicting Accuracy in the Subsample (n = 133) ...47
Table 3.1 Randomized Experiment: Pretest Assessment ........................................86
Table 3.2 Randomized Experiment: Posttest, Opposite Format, and Other Formats ....88
Table 3.3 Randomized Experiment: Intervention Problems .....................................89
I would not have completed this journey without the help of many people. First, my graduate advisor Nicole McNeil, who has guided me through six years of graduate school and supported my development as a scholar, writer, and presenter. Second, thanks to my quantitative minor advisor Lijuan (Peggy) Wang, who has provided me the opportunity to expand my quantitative expertise and also served as a teaching mentor. I am grateful for the valuable feedback from Nicole and Peggy as well as from the other members of my dissertation committee, Jill Lany and Sidney D’Mello. Thanks also to Mary Wagner Fuhs, for her support as a research collaborator and mentor.

This work would not have been possible without the support of many funding agencies. Individual studies examined in Study 1 were supported by funding I received from the National Science Foundation Graduate Research Fellowship DGE-1313583 and a Psi Chi Graduate Student Research Grant, Notre Dame Institute for Scholarship in the Liberal Arts Grants to myself and Casey Hall, a National Science Foundation CAREER Grant DRL-1054467 to Nicole McNeil, and Institute of Education Sciences Grants R305A110198 and R305B070297 to Nicole McNeil. Note that the opinions expressed here are my own and do not represent views of the National Science Foundation or the U.S. Department of Education. In addition, thank you to Susan Wagner Cook, Emily Fyfe, and Shanta Hattikudur for their willingness to share raw data for Study 1. Study 2 was supported by a Society for Research in Child Development Dissertation Funding
Award. Of course, this research would not be possible without the support of participating students, parents, teachers, and staff.

Thank you to current and former members of the Cognition, Learning, and Development Lab at Notre Dame for your assistance with experimenting, coding, and administrative tasks. Thanks especially to Lori Petersen, for your mentorship and friendship during my first two years of graduate school, and Connor O’Rear, for your companionship over the past three years as both a listener and collaborator. Thanks also to Bri Devlin, Cristina Carrazza, Mary McKeever, and Lauren Crawford for being there for me to bounce ideas off of and laugh with down in the basement of Haggar Hall.

Finally, I would like to thank my family and friends for their support, both throughout my time in graduate school and as long as I have known them. Thanks especially to my parents, Rob and Shirley, my husband, Dave, my in-laws, Jerry and Kristin, my sister, Kat, and my sister-in-law, Jess, and her husband, Kevin. Thanks to members of my extended family as well, especially the support of my grandmothers, Joy Byrd and Joy Pappas (1927-2012). The unconditional love and encouragement you all have given me have been instrumental in this journey. I also would like to thank my church family at River Valley Community Church. I would not be the person I am today without your influence.
Parents and teachers want to know how to structure the learning environment to optimize children’s learning. What they may not realize is that even relatively small and specific changes to the structure of the learning environment can have large effects on children’s understanding of important concepts, especially in mathematics. The knowledge children construct depends heavily on how educators present the material, from the specific way spaces are lined up in a numerical board game (Siegler & Ramani, 2009), to the types of gestures used during a lesson (Cook, Duffy, & Fenn, 2013), to the exact phrasing used in a word problem (Riggs, Alibali, & Kalish, 2015), to the specific letters used when writing variables in an algebraic equation (McNeil et al., 2010).

This dissertation examines how a seemingly small difference in presentation format affects children’s formal understanding of mathematical equivalence, the concept that two quantities are equal and interchangeable, symbolized by the equal sign (Kieran, 1981). Unfortunately, most elementary school children in the United States (ages 7-11) struggle to understand math equivalence (Baroody & Ginsburg, 1983; McNeil, 2007; Rittle-Johnson & Alibali, 1999). Children’s difficulties are shown in their solutions to math equivalence problems, problems that have operations on both sides of the equal sign (e.g., $3 + 4 = \_ \_ + 5$), as most children either add only the numbers before the equal sign (e.g., $3 + 4 = 7 + 5$) or add up all of the numbers in the problem (e.g., $3 + 4 = 12 + 5$);

It is important to determine the best methods for improving children’s understanding of math equivalence because it has been identified as a foundational concept in mathematics that, once learned, can make it easier to learn new mathematical concepts (Charles, 2005; National Council of Teachers of Mathematics, [NCTM], 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSSO], 2010). Indeed, as early as first grade, one of the new Common Core Standards (Standard 1.OA.D.7) is to “understand the meaning of the equal sign, and determine if equations involving addition and subtraction are true or false.” Furthermore, it is widely assumed that math equivalence is one of the most important concepts for developing algebraic thinking (Alibali, Knuth, Hattikudur, McNeil, & Stephens, 2007; Blanton & Kaput, 2005; Carpenter, Franke, & Levi, 2003; Falkner, Levi, & Carpenter, 1999; Knuth, Stephens, McNeil, & Alibali, 2006; Stephens, Blanton, Knuth, Isler, & Gardiner, 2015). Unfortunately, many students in the U.S. struggle to understand algebraic concepts and procedures (Knuth et al., 2006; MacGregor & Stacey, 1997; Sfard, 1991), and first-year algebra courses have even been called an “unmitigated disaster” for most students (National Research Council [NRC], 1998, p.1). This is concerning because algebra has been labeled as a “gatekeeper” to future educational and employment opportunities (Moses & Cobb, 2001; NRC, 1998), and some
experts have suggested that difficulties with algebra may be a major reason for high school and college dropout (Hacker, 2012). A formal understanding of math equivalence is not only predictive of performance solving algebraic equations (Knuth et al., 2006) and use of advanced algebraic reasoning (Alibali et al., 2007), but it predicts future math achievement, even after controlling for IQ, SES, gender, and baseline math achievement (McNeil, Hornburg, Devlin, Carrazza, & McKeever, 2017). Given its importance, it is critical for educators to determine how to best teach this fundamental concept.

One aspect of the environment that may matter when children are learning math equivalence is the specific format in which problems are written (i.e., the exact positions of the numbers and symbols). Research on math equivalence has primarily used two formats: “left-blank” problems, which have the blank immediately after the equal sign, on the left of the expression after the equal sign (e.g., $3 + 4 = \_\_ + 5$), and “right-blank” problems, which have the blank at the end, on the right of the expression after the equal sign (e.g., $3 + 4 = 5 + \_\_$). Thus, when teaching a lesson on math equivalence, a teacher could choose to write either type of problem on the board. This distinction may not seem important because the problems are the same conceptually and the computation required for solving is the same; however, research in other areas of mathematics indicate the format in which problems are presented matters for performance and learning. For example, the numerical surface form (e.g., “3 + 4 = ?” vs. “three + four = ?”) affects arithmetic calculation efficiency (Campbell, 1994) and strategy use (Campbell & Fugelsang, 2001). Problem orientation (horizontal vs. vertical) affects solving of multidigit arithmetic problems (Trbovich & LeFevre, 2003). Even the side on which the operation is presented in simple addition problems (e.g., $3 + 4 = \_\_ vs. \_\_ = 3 + 4$) affects
children’s understanding of the underlying concepts (McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011). Thus, small changes to the positions of numbers and symbols in math equivalence problems may affect children’s understanding of math equivalence.

My objective for this dissertation was to determine how the specific format in which math equivalence problems are written (i.e., the positions of the numbers and symbols) affects performance and learning. Surprisingly, no research to date has directly examined the effect of problem format on children’s problem-solving performance or learning. In terms of performance, research in the domain of math equivalence has largely collapsed performance across problem types (as long as problems had operations on both sides), and reported overall accuracy. The nature of intervention studies (aimed at teaching children who cannot solve any math equivalence problems correctly at a pretest) excludes children who do solve some problems correctly; thus, unfortunately, this subpopulation has largely gone unstudied. One exception is a paper by Rittle-Johnson, Matthews, Taylor, and McEldoon (2011), which reported the item difficulty (calculated from average accuracy) of each math equivalence problem used in two forms of their assessment. Accuracy on the left-blank item was significantly lower than accuracy on the right-blank item. However, this study alone is not sufficient to state that left-blank problems are more difficult to solve correctly than right-blank problems because the study was not designed to test that question. The only significant difference emerged between performance on a single left-blank problem (4 + 5 + 8 = __ + 8) and a single right-blank problem (7 + 6 + 4 = 7 + __). Thus, it is unclear if it is the blank position or the numbers involved in the problems that make one easier than the other. Moreover, both problems included three addends on the left side and a repeated addend on the right.
side (i.e., a number from the left side was repeated as the addend on the right side). Other types of math equivalence problems may not show this difference. Furthermore, even if the findings from Rittle-Johnson et al. (2011) could be taken as evidence that left-blank problems are more difficult to solve correctly than right-blank problems, it would remain unclear why left-blank problems may be more difficult. It is important to address this gap in the literature to provide a critical test of one of the prevailing theories of children’s learning across domains, that poor encoding causes solving difficulties.

1.1 Links Between Problem Encoding and Problem Solving

One prevailing explanation for children’s poor performance solving math equivalence problems is their inaccurate encoding of the problems. Across many domains, changes in the encoding, or representation of problems, is widely cited as a mechanism driving cognitive development and allowing children to overcome their misconceptions. There is a long line of evidence demonstrating that with age and experience there are improvements in encoding of problems (e.g., Booth & Davenport, 2013; Chase & Simon, 1973; Siegler, 1976; Sternberg, 1977). In the domain of decimal fractions, Rittle-Johnson and colleagues argued that problem representation was the mechanism underlying iterative relations between conceptual and procedural knowledge (Rittle-Johnson, Siegler, & Alibali, 2001). Encoding accuracy is also linked to accuracy solving math equivalence problems (Alibali, McNeil, & Perrott, 1998; McNeil & Alibali, 2005b), balance scale problems (Siegler, 1976), and Tower of Hanoi problems (Kotovsky, Hayes, & Simon, 1985).

Based on this account, problems that are encoded more accurately will both be easier to solve correctly and better for instruction. This is because when children attempt
to solve problems that are encoded accurately, they may have more cognitive resources to
generate new problem-solving strategies, and when learning with problems that are
encoded accurately, children are able to allocate their resources to learning correct
conceptual information about problem features rather than spending resources on
encoding the exact features. Crooks and Alibali (2013) provided some support for this
view with undergraduates, showing that poor problem encoding mediated the relation
between activation of traditional arithmetic patterns and poor solving of math
equivalence problems.

Previous research has found that children’s encoding of math equivalence
problems depends on format, with encoding accuracy on right-blank problems (e.g., $a + b$
$= c + ___) being much lower than encoding accuracy on left-blank problems (e.g., $a + b =$
___ + c; McNeil & Alibali, 2004). This is because right-blank problems are more
perceptually similar to the traditional format with the blank at the end ($a + b + c = ___$)
and are more likely than left-blank problems to be incorrectly internally represented (i.e.,
misencoded) as traditional arithmetic problems (e.g., $3 + 4 = 5 + ___$ is more likely than $3$
$+ 4 = ___ + 5$ to be misencoded as $3 + 4 + 5 = ___$). In fact, McNeil and Alibali (2004)
found that children were just as accurate at encoding left-blank problems (e.g., $a + b + c$
$= ___ + d$) as they were at encoding traditional problems (e.g., $a + b + c = ___$ or $a + b + c$
$+ d = ___$). Following this account, if poor encoding causes solving difficulties, then right-
blank problems should be harder to solve correctly and learn from than left-blank
problems.
1.2 The Change-Resistance Account

Another related theory of children’s difficulties is the change-resistance account (McNeil & Alibali, 2005b), which, when accompanied by the aforementioned encoding-solving account, makes the same prediction that right-blank problems will be harder to solve correctly and learn from than left-blank problems. McNeil and Alibali’s (2005b) change-resistance account argues that children’s difficulties with math equivalence are due to overly narrow experience with arithmetic in the early school years (see also Baroody & Ginsburg, 1983; McNeil, 2014; McNeil, Hornburg, Fuhs, & O’Rear, 2017).

In the U.S., children’s arithmetic learning has a strong procedural focus, without much explicit attention to the equal sign. Most often, the goal is to solve for the “answer,” and problems are almost always presented with operations to the left of the equal sign and the “answer blank” at the end (e.g., $3 + 4 = _ _$; McNeil et al., 2006; Powell, 2012; Seo & Ginsburg, 2003), a format that fails to emphasize that the two sides of an equation are interchangeable. According to the change-resistance account, children detect and extract the patterns routinely encountered in traditional arithmetic to the point of developing representations in long-term memory that become defaults when they encounter novel math problems.

Children learn to expect math problems to be set up with all operations on the left side of the equal sign in an “operations equals answer” format (Alibali, Phillips, & Fischer, 2009; Cobb, 1987; McNeil & Alibali, 2004). Children learn to interpret the equal sign as a “do something” symbol (Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols, 1980; Herscovics & Kieran, 1980; McNeil & Alibali, 2005a). Finally, children learn to “perform all given operations on all given numbers” as their go-to strategy when
solving math problems (McNeil, 2007; McNeil & Alibali, 2005b; Rittle-Johnson, 2006). These three patterns become entrenched over time and children rely on them as the default when encoding, interpreting, and solving mathematics problems. Reliance on those patterns leads to success on traditional arithmetic problems (e.g., $3 + 4 = \_\_\_\_$), but errors emerge on math equivalence problems because children’s representations overlap with the novel problems but do not map exactly onto them (cf. Bruner, 1957; Zevin & Seidenberg, 2002).

A key assumption of the change-resistance account is that activities and lessons that activate children’s narrow view of arithmetic will hinder children’s understanding and learning of math equivalence (Crooks & Alibali, 2013; McNeil, 2008; McNeil & Alibali, 2004; McNeil & Alibali, 2005b; McNeil, Rittle-Johnson, Hattikudur, & Petersen, 2010). One specific prediction that follows logically from this assumption is that it will be more difficult for children to solve correctly and learn from right-blank problems than leftblank problems, because right-blank problems activate children’s traditional arithmetic schema and are more likely than left-blank problems to be misencoded as traditional arithmetic problems; McNeil & Alibali, 2004).

Researchers and educators are eager to determine the best methods for helping children overcome their difficulties with math equivalence, and many have relied on the theory that poor encoding causes solving difficulties and the subsequent prediction paired with McNeil and Alibali’s (2005b) change-resistance account when developing new interventions (e.g., McNeil, Fyfe, & Dunwiddie, 2015; Powell, Driver, & Julian, 2015). Specifically, the view that left-blank problems will be easier to solve correctly and learn from because they are encoded more accurately is widely held by researchers who study
children’s understanding of math equivalence, with several researchers only using left-blank problems in interventions that teach children how to solve math equivalence problems (e.g., Alibali, 1999; Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014; Siegler, 2002). However, this prediction has never been tested directly, and at least two alternative accounts in the literature suggest that left-blank problems will be more difficult to solve correctly and learn from than right-blank problems.

1.3 Alternative Accounts of Children’s Performance

Following the prevailing account, if poor encoding causes solving difficulties, then right-blank problems should be harder to solve correctly than left-blank problems. However, this may not be the case. A second, equally plausible prediction, is that the role of encoding in solving differs based on problem format. It is possible that children can encode features correctly but still lack conceptual understanding of the problem structure. In line with this framework, Booth and Davenport (2013) found that although both accurate encoding of problem features and conceptual understanding of problem features were correlated with equation-solving success, only conceptual understanding of the features was independently predictive of success. Thus, how children interpret particular problems may be more important than their overall accuracy remembering the problem features. From this view, encoding accuracy may matter more for some types of problems than others, and above all, it may not matter if you encode the problem perfectly if you don’t understand what the numbers and symbols in the problem mean.
1.3.1 Run-on Schema

In line with this view, left-blank math equivalence problems may be more
difficult to solve correctly than right-blank problems because even when a left-blank
problem (e.g., $3 + 4 = \_ + 5$) is encoded accurately, it can still be misconceptualized as a
“run-on” traditional addition problem with two steps (e.g., $3 + 4 = 7 + 5$ because $3 + 4 = 7$ and then $7 + 5 = 12$). In this way, children’s traditional arithmetic schema leads them to
incorrectly interpret the meaning of the structure they encoded. Left-blank problems (e.g.,
$3 + 4 = \_ + 5$) can be interpreted as run-on problems; however, it is more difficult to
misconceptualize a right-blank problem such as $3 + 4 = 5 + \_$ as a run-on problem
because $3 + 4$ does not equal $5$.

Following the change-resistance account (McNeil & Alibali, 2005b), this
activation of a traditional arithmetic schema when left-blank problems are encoded
accurately leads to misconceptualization of left-blank problems that will hinder success.
Thus, it is important to note that this run-on schema is simply a different hypothesis
stemming from ideas of change-resistance without assuming that encoding performance
predicts solving performance, rather than a hypothesis that directly conflicts with the
change-resistance account.

It is not uncommon for teachers to reinforce this run-on schema by using the
equal sign incorrectly “to represent a string of calculations” (e.g., $20 + 30 = 50 + 7 = 57 + 8 = 65$) (Carpenter et al., 2003). Some teachers also use “=” incorrectly to demonstrate
partitioning strategies (e.g., $45 + 22 = 40 + 20 = 60 + 5 = 65 + 2 = 67$). In a guide for
teachers illuminating children’s misconceptions about the equal sign, one teacher
recognized she had “regularly used the equal sign when it is not mathematically
appropriate” (Cockburn & Littler, 2008, p. 30). Teachers may be unintentionally saying or writing things in math lessons that conform to this run-on schema, and children are not likely to be presented with any information that would counteract their misconception, as the majority of problems presented are in the “operations equals answer” form (McNeil et al., 2006; Powell, 2012; Seo & Ginsburg, 2003). Calculators also reinforce the run-on schema when children perform calculations, press the “=” key, perform more operations, and then press the “=” again at the end to obtain the answer to the programmed sequence (Hughes, 1986).

1.3.2 Difficult Unknown

Another possible reason to predict left-blank problems to be more difficult to solve correctly than right-blank problems is because the unknown in a left-blank problem is more difficult; specifically, it is a “start unknown” (\( _\_ + c \)), whereas the unknown in a right-blank problem is a “change unknown” (\( c + _\_ \)). For traditional addition problems, start unknowns are more difficult than change unknowns (e.g., \( _\_ + 5 = 7 \) is more difficult than \( 5 + _\_ = 7 \), Carpenter, Fennema, Peterson, & Carey, 1988; Clements & Sarama, 2014). When the initial quantity is unknown, it cannot be represented directly with objects (e.g., fingers on a hand), but when the change is unknown, children can easily count up from the given number to find the unknown. It is, therefore, possible that the placement of the unknown has similar effects for math equivalence problems as it does for traditional addition problems as solvers hold the sum of the left side of the equation in mind while attempting to determine the unknown number on the right side of the equation (in either the “start” or “change” position). It may be easier to have a
number on the right side of the equation in the first position and then count up from that number to reach the equivalent sum (as would be the case for a right-blank problem).

1.4 The Role of Encoding in Children’s Performance

In this dissertation, I tested key predictions about the role of encoding in children’s solving of math equivalence problems by determining which problem format is more difficult for children to solve correctly (left-blank or right-blank). Furthermore, if results did not support the prevailing account of encoding performance predicting solving performance (and one interpretation of the change-resistance account, that children’s traditional arithmetic schema is activated when right-blank problems are poorly encoded), then I could distinguish between the run-on schema and difficult unknown accounts with a closer examination of the data. The prevailing account that encoding performance predicts solving performance predicts that right-blank problems \((a + b = c + \_\_)_1\) are more difficult to solve correctly than left-blank problems \((a + b = \_\_ + c)_1\). However, the alternative accounts (discussed above) predict that left-blank problems are more difficult for children to solve correctly than right-blank problems. Specifically, if left-blank problems are more difficult than right-blank problems because they are easily misconceptualized as run-on equations, then individual differences in encoding should not be highly predictive of accuracy on left-blank problems (because accurately encoded left-blank problems can still be interpreted as run-on equations). Moreover, children who solve the left-blank problems incorrectly should be more likely to solve them with the add-to-equal-sign strategy (e.g., \(3 + 4 = 7 + 5\)) than other incorrect strategies. In contrast, if the difficulty of the unknown position underlies the increased difficulty of left-blank problems relative to right-blank problems, then children should be more likely to make
arithmetic mistakes and be one or two away from the correct solution when they explain that they used a correct strategy (e.g., $3 + 4 = 1 + 5$, explaining “because three plus four is six and one plus five is six”) on left-blank problems than on right-blank problems.

1.5 The Role of Encoding in Children’s Learning

Along with determining which math equivalence problem format is easiest for children to solve correctly, my second aim was to determine which math equivalence problem format is best to use when instructing children on how to solve the problems. The difference in children’s encoding accuracy between the two problem formats provides a straightforward opportunity to test competing predictions about the role of encoding in learning. Some theories suggest that using problems that are easier to encode will facilitate learning. This seems intuitive, as students and teachers tend to measure the success of a lesson based on the ease of encoding information during instruction (Bjork, 1994). In other words, better performance during instruction should translate to better learning. More formally, cognitive load theory suggests that students’ learning will be optimized when instruction is designed to reduce extraneous cognitive load (Sweller, 1988, 1994). This is because students who are burdened by extraneous cognitive load during learning cannot devote all of their cognitive resources to the learning process and, thus, are not able to construct the depth of knowledge that other students may be able to construct (Sweller, van Merriënboer, & Paas, 1998). If students are devoting all of their resources to encoding, there may not be resources left over for conceptual understanding or strategy generation (Alibali, Crooks, & McNeil, 2017).

Similarly, one interpretation of the change-resistance account (McNeil & Alibali, 2005b) also assumes that right-blank problems (e.g., $a + b = c + _$) are less effective as
instructional tools when teaching about math equivalence than left-blank problems (e.g., \(a + b = \_ \_ + c\)). This follows from the thought that right-blank problems are more likely to activate children’s knowledge of traditional arithmetic as a result of children’s tendency to misencode them as traditional addition problems (McNeil & Alibali, 2004). According to McNeil (2008) “children should have more difficulty learning mathematical equivalence when their narrow knowledge [of traditional arithmetic] is activated than when it is not activated” (p. 1533). In support of this view, McNeil (2008) found that children who were taught about the concept of equivalence in the context of traditional arithmetic problems (e.g., \(13 + 15 = 28\)) performed worse on math equivalence problems than children who were taught in the context of nontraditional problems (e.g., \(28 = 28\)). Thus, if activation of traditional arithmetic knowledge is considered based on encoding errors, right-blank problems should activate that narrow knowledge and hinder learning.

However, it does not necessarily follow that whatever format is encoded more inaccurately is also worse for using during instruction on how to solve the problems. In fact, the notion of desirable difficulties (Bjork, 1994; see also Bjork & Bjork, 2011, McDaniel & Butler, 2010) suggests the opposite. On a typical measure of encoding in which children reconstruct problems after viewing them for five seconds each (cf. Chase & Simon, 1973; Siegler, 1976), children may experience the encoding event easily for both format types but actually misencode right-blank problems (McNeil & Alibali, 2004). However, the encoding experience may be quite difficult in the context of an intervention, in which children receive feedback that they are incorrectly solving the problem and explicit direction as to the location of the equal sign in the problem.
Problems that invoke a difficult encoding experience may in fact be more
effective platforms for instruction because they may lead to deeper processing during
instruction, and lead to improved retention of the material. For example, Benjamin and
colleagues demonstrated that items answered with the most difficulty during an initial test
were found to be the most memorable at a later test, counter to participants’ predictions
(Benjamin, Bjork, & Schwartz, 1998). In this way, the difficulty searching for a solution
creates a more elaborated encoding event in episodic memory, which can then be
accessed more easily later on a free-recall task (Gardiner, Craik, & Bleasdale, 1973).

Although counter-intuitive, in many cases learning phases that are more effortful
actually can lead to more durable and flexible learning (e.g., Diemand-Yauman,
Oppenheimer, & Vaughan, 2011; Richland, Bjork, Finley, & Linn, 2005). Thus, even
though children are much less accurate at encoding right-blank problems than left-blank
problems, right-blank problems could still be the better format to use when teaching
children how to solve the problems. In fact, children’s tendency to misencode right-blank
problems as traditional addition problems may be an advantage because it provides a
straightforward opportunity to prompt cognitive conflict (Inhelder, Sinclair, & Bovet,
1974; Limón, 2001; Lucariello, 2009; Piaget, 1980; VanLehn, 1996). Consistent with this
view, several previous studies have involved interventions designed to direct children’s
attention to the position of the equal sign (Alibali et al., 2017; Alibali et al., 1998; McNeil
& Alibali, 2005b). When interventions improve children’s encoding of important
problem features in this way, improvements in conceptual understanding and solving
accuracy should follow (Rittle-Johnson et al., 2001). It may generally be better to use
problems that are difficult to correctly encode during instruction because it leads to
deeper processing of the problems, facilitating encoding and retrieval (Craik & Tulving, 1975; Diemand-Yauman et al., 2011).

The second study of my dissertation provides a critical test of these competing accounts, and also contributes to the development of interventions aimed at improving children’s understanding of math equivalence, which is one of the foundational concepts in mathematics that, once learned, makes it easier for children to acquire new math concepts (Charles, 2005). If problems that are often misencoded hinder learning, then one possible reason why right-blank problems may be worse for instruction than left-blank problems is because children are much less accurate when encoding right-blank problems, and thus, may be burdened with extraneous cognitive load when given instruction on those problems (Sweller, 1988; Sweller et al., 1998). However, if problems that are often misencoded facilitate learning, then one possible reason why right-blank problems may be better for instruction is because children’s tendency to misencode right-blank problems can provide an opportunity for deeper processing of the material (Craik & Tulving, 1975; Diemand-Yauman et al., 2011).
2.1 Method

2.1.1 Design

I conducted an integrative data analysis of children’s performance on pretest assessments from 14 previously conducted studies. In single studies in the literature, typically only about 15%-25% of children are able to solve at least one problem correctly (e.g., Alibali, 1999; Alibali et al., 1998; McNeil & Alibali, 2005b). Integrative data analysis allowed me to test my hypotheses with greater power, because the phenomenon of interest (i.e., accuracy on each problem type) occurs at quite a low base rate (Curran & Hussong, 2009). With a greater overall sample size, I also gained precision and a less biased estimate of the effect. This analysis was part of a larger quantitative project in fulfillment of the Minor in Quantitative Psychology at the University of Notre Dame in which I compared the procedures, results, and conclusions from meta-analysis and integrative data analysis (Hornburg, 2016; Hornburg, Wang, & McNeil, 2017).

The 14 studies were chosen because they assessed children’s understanding of math equivalence using the same tasks, including both right-blank and left-blank problems, and were either conducted, published, or under review between January 2010 and July 2016. I restricted inclusion to studies in which the pretest did not include
traditional arithmetic problems (e.g., $a + b + c = \_\_\_$) or nontraditional arithmetic problems (e.g., $a = b + \_\_\_$) intermixed with math equivalence problems, in order to understand differences between performance on problems with operations on both sides without any outside influence of seeing these other problem formats. Ten of the 14 studies also included an encoding assessment. Study characteristics including sample size, data collection period, assessment method, grade and gender distribution, as well as session location and season are displayed in Table 2.1, ordered by year of data collection.

2.1.2 Participants

From 14 studies, data for problem solving was compiled from a total of 1,414 children (731 girls, 681 boys, gender was not reported for two children). Across the 14 studies, there were 18 children who participated in two separate studies. The data point at which the most information about the participant was obtained (i.e., problem solving and problem encoding) was kept in the full dataset ($N = 1,414$), and the duplicate participant data was excluded. However, results are unchanged if these 18 children are excluded altogether.

Children’s grade level ranged from first to fifth, with the majority of children (87%) in either second or third grade ($M = 2.64$, $SD = 0.73$; one child’s grade was not reported). Children assessed during the summer were labeled as the grade they would be entering. Children’s age ($N$ reported = 890) ranged from 6 years 11 months to 11 years 6 months ($M = 8$ years 7 months, $SD = 10$ months). Most children ($N$ reported = 883) were between the ages of 7 and 10; four children were not yet 7 years old and three children were already 11 years old (conclusions remain unchanged when these children are...
The racial/ethnic makeup of the sample ($N$ reported = 650) was 65% white, 17% African American or black, 10% Hispanic or Latino, 3% Asian, and 5% other.

Encoding data was compiled from ten studies, from a total of 984 children (504 girls, 480 boys). Children’s grade level ranged from first to fifth ($M = 2.64$, $SD = 0.80$; one child’s grade was not reported). Children assessed during the summer were labeled as the grade they would be entering. Children’s age ($N$ reported = 692) ranged from 6 years 11 months to 11 years 6 months ($M = 8$ years 8 months, $SD = 10$ months). The racial/ethnic makeup of the sample ($N$ reported = 452) was 65% white, 17% African American or black, 12% Hispanic or Latino, 3% Asian, and 5% other.

2.1.3 Measures

2.1.3.1 Problem Solving

In all 14 studies, children solved a set of math equivalence problems including both left-blank formats (e.g., $1 + 5 = \_ \_ + 2$) and right-blank formats (e.g., $2 + 7 = 6 + \_ \_ \_ \_ \$). All problems had either two or three addends on the left side of the equal sign and an addend and a blank on the right side (either in the left-blank or right-blank format), though the exact numbers varied. The majority of children (93%; 11 studies) solved problems in a set order; however, in Studies 3, 4, and 5 children were randomly assigned to solve either a left-blank problem first or a right-blank problem first. The specific problems and particular order given in each study are detailed in Table 2.2.

In each study, children were told to try their best to solve the problems and to write the number that goes in the blank. The majority of children (61%; 11 studies) were then asked how they got their solution (cf. Alibali, 1999; Rittle-Johnson & Alibali, 1999).
Across the studies, children solved between two and eight problems ($M = 4.34$, $SD = 1.81$). Half of the children (50%; nine studies) solved four problems. Although the total number of math equivalence problems varied, all children solved an equal number of left-blank compared to right-blank problems. Because children did not solve the same number of problems, data were analyzed in terms of accuracy. As noted above, accuracy was considered only on math equivalence problems solved prior to children solving any other type of problem (e.g., $a + b + c + d = _-_ _$).

Correctness was determined from the number the child wrote in the blank. Strategies were coded using a system from previous research (e.g., McNeil, 2007; McNeil & Alibali, 2004; Perry et al., 1988). For all studies, the problem-solving strategy was first coded by what the solution alone indicated. Responses were coded as reflecting a particular strategy as long as they were within ±1 of the response that would be achieved with that particular strategy.¹ If the solution in the blank was unclear and verbal explanations were collected, then the child’s verbal strategy was considered. For studies in which verbal explanations were not collected, solutions that were within ±1 of the correct solution were coded as correct only if the child had solved another problem exactly correct. Table 2.3 presents common strategies used to solve math equivalence problems with example verbal explanations.

¹ One exception to this rule was that for Study 6 (see Table 2.1) children’s solutions could not be obtained from the authors, only correctness, which was determined with the strict system of exactly correct or not.
2.1.3.2 Problem Encoding

In ten of the 14 studies, children completed an encoding task after the solving task. Children were shown a series of math equivalence problems for five seconds each, and after each problem was put down children wrote exactly what they saw (cf. Chase & Simon, 1973; Siegler, 1976). Across the ten studies, children were given between two and four problems to encode (note that only performance encoding math equivalence problems prior to encoding any traditional problems was considered). Most children (78%; seven studies) encoded four problems. In all but one study (Study 2) children encoded both left-blank and right-blank problems. The specific problems and particular order given in each study are detailed in Table 2.2. Similar to problem solving, data were analyzed in terms of conceptual accuracy (described in more detail below) because children were not given the same number of problems to encode.

Children’s responses were coded based on a system from previous research (e.g., McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999). This coding system distinguishes between number errors, such as mixing up the order of the numbers, and conceptual errors, such as omitting the equal sign, omitting the plus on the right side of the equal sign, or converting the problem to a traditional arithmetic problem.

2.2 Data Analytic Plan

I used multilevel modeling to account for the nested data structure where participants are nested within study, and as such, there is a lack of statistical independence among participants (i.e., participants from a particular study may be more alike in their responses than participants from different studies). Multilevel models are often used in educational data where students are nested in classrooms or classrooms are
nested within schools (Bock, 1989; Raundenbush & Bryk, 1986). All available data were
used in analyses. I conducted the integrative data analyses using the PROC MIXED
procedure in SAS (v9.4). Before using that procedure, I calculated Cohen’s $d$
standardized mean difference (right-blank accuracy – left-blank accuracy) for each
individual using $S_{within}$ for the study (denoted $i$) in which the individual (denoted $j$) was
nested (see equations below). Akin to a repeated measures analysis, the correlation
between accuracy on each problem type was incorporated into the calculation. I decided
to have the $d$ outcome right-blank accuracy – left-blank accuracy so that a positive $d$
would be in the hypothesized direction, in line with the alternative view that conceptual
understanding of problem features matters more than accurate encoding. Equations used
to calculate $d$ are displayed below.

$$d_{ij} = \frac{y_{1ij} - y_{2ij}}{S_{i, within}}$$

$$S_{i, within} = \frac{S_{i, diff}}{\sqrt{2(1 - r_i)}}$$

2.2.1 Unconditional Models

I set the degrees of freedom method to Kenward-Roger for the unconditional
models, because of the small number of studies in which children were nested. Equations
for the unconditional models are displayed below.

$$d_{ij} = \beta_{0i} + e_{ij}$$

$$\beta_{0i} = \beta_{00} + U_{0i}$$

where $d_{ij}$ is the observed effect size for study $i$, $\beta_{0i}$ is the population effect size for study
$i$, and $\beta_{00}$ is the population overall effect size.
Additionally, the variance of $U_{0i}$ is equal to $\tau^2$. Note that in the unconditional models and the conditional models below the random effects were expected to be normally distributed.

2.2.2 Conditional Models

I set the degrees of freedom method to Between-Within for the various conditional models, because I wanted to evaluate both between-study and within-study effects. For the conditional models, I created centered variables for the number of problems solved (around the grand mean) and number of problems encoded (around the grand mean) as well as grade, gender, and encoding accuracy (around each group mean), in order to disaggregate between-study and within-study effects. I conducted individual models with method (paper and pencil or at a whiteboard) and number of problems solved as study-level variables. For the remaining individual-level predictors (grade, gender, and encoding accuracy), I included both the study-level average and the individual-level predictor in each model. In the model with encoding accuracy as a predictor, the number of problems encoded was included as a control variable. The within-study random effect was estimated in each model. Equations for the conditional models with a study-level predictor are displayed below.

$$d_{ij} = \beta_{0t} + e_{ij}$$

$$\beta_{0t} = \beta_{00} + \beta_{01}X_t + U_{0t}$$

where $\beta_{01}$ is the between-study effect.

Similarly, equations for the conditional models with an individual-level predictor are displayed below.

$$d_{ij} = \beta_{0i} + \beta_{1i}(X_{ij} - \bar{X}_r) + e_{ij}$$
\[
\beta_{0i} = \beta_{00} + \beta_{01} x_i + u_{0i} \\
\beta_{1i} = \beta_{10} + u_{1i}
\]

where \( \beta_{01} \) is the between-study effect and \( \beta_{10} \) is the within-study effect.

2.3 Results

2.3.1 Overall Performance

Prior to reporting results from the multilevel modeling analyses, I first present some overall descriptive statistics. Overall problem solving performance was poor, consistent with previous literature (e.g., Alibali, 1999; Alibali et al., 1998, McNeil & Alibali, 2005b). The majority of children (74%) solved zero problems correctly. Only 12% of children demonstrated 100% accuracy across all problems. The distribution of children’s accuracy on both left-blank and right-blank problems is displayed in Figure 2.1. Table 2.4 displays problem solving results for each study separately.

On the problem encoding assessment, 34% of children did not encode any problems correctly, and few (24%) encoded all problems without any conceptual errors. Table 2.5 displays problem encoding results for each study separately. The distribution of children’s overall conceptual encoding accuracy is displayed in Figure 2.3.

2.3.2 Unconditional Models of Performance by Problem Format

My unconditional multilevel model yielded an effect size \( d = 0.14 \) (\( SE = 0.03 \)), \( p < .001 \), 95% CI [0.07, 0.21], observed power = .98. This \( d \) corresponded to significantly greater accuracy on right-blank problems (21%) compared to left-blank problems (16%) in this full sample. Note that because one of the 14 studies (Study 6) did not have exact
solutions available for analysis and used a slightly stricter coding scheme (see Method section above), analyses were conducted excluding this study (13 studies $N = 1,182$), and results remained unchanged, $d = 0.16$ ($SE = 0.03$), $p < .001$, 95% CI [0.10, 0.22], observed power > .99, thus, Study 6 was kept in for further analyses. Because a disproportionate percentage (89%) of children in the full sample ($N = 1,414$) solved a left-blank problem first, I conducted analyses separately by which type children saw first. The pattern of results held no matter which format children solved first (19% vs. 15% for the subsample of children who solved a left-blank problem first ($N = 1,252$), $d = 0.14$ ($SE = 0.04$), $p < .001$, 95% CI [0.07, 0.22], and 31% vs. 26% for the subsample of children who solved a right-blank problem first ($N = 162$), $d = 0.11$ ($SE = 0.05$), $p = .02$, 95% CI [0.01, 0.21]).

Finally, similar to prior research with children’s performance solving math equivalence problems in which performance is examined based on whether children solve at least one problem correctly or solve at least 75% of problems correctly because overall performance is so skewed (e.g., Byrd et al., 2015; McNeil, 2007), I tested whether there was a significant difference between the proportion of children who were able to solve at least one right-blank problem correctly (259 out of 1,414) and the proportion of children who were able to solve at least one left-blank problem correctly (337 out of 1,414). These proportions were significantly different, $z = 3.59$, $p < .001$, in that a significantly greater proportion of children were able to solve at least one right-blank problem correctly than at least one left-blank problem correctly. In addition, the proportion of children who were able to solve at least 75% of right-blank problems correctly (260 out of 1,414) was
significantly greater than the proportion of children who were able to solve at least 75% of left-blank problems correctly (203 out of 1,414), $z = 2.80, p = .005$.

Furthermore, when the sample is restricted to children who solved at least one but not all math equivalence problems correctly ($n = 196$), the same pattern emerged in which children were more accurate solving right-blank problems, $d = 1.03 (SE = 0.23), p < .001, 95\% \text{ CI} [0.58, 1.48]$. Note that $d$ for each individual in this analysis was calculated using the standard deviation within each study just as before (see Equations above). The overall effect size $d$ corresponded to significantly greater accuracy on right-blank problems (60%) compared to left-blank problems (27%) for this subsample of children with emerging knowledge of math equivalence. The distribution of children’s accuracy on both left-blank and right-blank problems for this subsample is displayed in Figure 2.2, and the distribution of children’s overall conceptual encoding accuracy for this subsample is displayed in Figure 2.3.

Again, because a disproportionate percentage (91%) of children solved a left-blank problem first, I conducted analyses separately by which type children saw first. The pattern of results held no matter which format children solved first (59% vs. 28% for the subsample of children who solved a left-blank problem first and 69% vs. 25% for the subsample of children who solved a right-blank problem first).

In addition, when the analysis was limited to the subsample of children who solved only one math equivalence problem correctly (and saw multiple problems in each format), more children solved only a right-blank problem correctly ($n = 41$) than only a left-blank problem correctly ($n = 20$), $p = .003$. 
Results from these unconditional models are contrary to the prevailing account that poor encoding causes solving difficulties. I next tested the significance of various potential moderators of this effect.

2.3.3 Conditional Models of Performance by Problem Format

Besides the unconditional model in which I tested the significance of the overall effect size accounting for participants nested within studies, I also continued the model building process by adding level-1 variables as fixed effects, then as random effects, and finally including level-2 variables as fixed effects in order to determine what variables had a significant impact on the standardized mean difference observed. Models tested included level-2 averages as control variables as well. Results from the best fitting model are summarized in Table 2.6. The best fitting model included level-1 predictors of grade, gender (0 = boy, 1 = girl), and encoding accuracy, level-2 predictors of assessment method (0 = paper, 1 = board) and number of encoding problems (as a control), and study (14 groups) as the level-2 covariate. In the full model ($N = 980$), no between-study coefficients were significant. The within-study coefficient for gender was significant, $\hat{\beta} = -0.10, p = .02$, as was the within-study coefficient for encoding accuracy, $\hat{\beta} = 0.14, p = .04$, but the within-study coefficient for grade was not. Controlling for the other variables, on average, girls had significantly smaller differences in problem-solving accuracy between formats (right blank – left blank), compared to boys in the same study. In addition, children who were more accurate encoders relative to other children in their particular study tended to have greater accuracy solving right-blank problems relative to left blank problems.
Next, I conducted several follow-up analyses to gain insight into whether the “run-on schema” or the “difficult unknown” account provides a better explanation for this result.

2.3.4 Relations Between Solving and Encoding

If children’s run-on schema was underlying the performance difference, then children may be easily encoding the positions of the numbers and symbols in the left-blank problems, but be misinterpreting them as run-ons. Therefore, the relation between encoding and solving may differ between the two problem formats. Evidence from the conditional model above indicated that encoding accuracy was a significant predictor of the standardized mean difference (right-blank accuracy – left-blank accuracy) in the overall sample, $\beta = .14, p = .04$.

I also examined the correlations between solving accuracy and encoding accuracy to address this in a way that is more interpretable. For the subsample of children who were able to solve at least one math equivalence problem correctly but not all correctly and who also had data on the encoding measure ($n = 133$), the correlation between overall encoding conceptual accuracy and accuracy solving right-blank problems was significant, $r = .27, p = .002$. However, the correlation between overall encoding conceptual accuracy and accuracy solving left-blank problems was not significant, $r = .09, p = .30$. When I separated encoding conceptual accuracy into encoding for left-blank problems and encoding for right-blank problems, the same pattern was displayed. The correlation between encoding conceptual accuracy for right-blank problems and accuracy solving right-blank problems ($n = 133$) was significant, $r = .23, p = .007$. However, the
correlation between encoding conceptual accuracy for left-blank problems and accuracy solving left-blank problems \((n = 121\); one study did not encode left-blank problems) was not significant, \(r = .12, p = .19\).

In addition to analyzing correlations with this subsample \((n = 133\), I also used the multilevel modeling framework accounting for participants nested within study to test the significance of encoding accuracy as a predictor of performance on each problem type while controlling for other important variables. The same variables were included in each model as were used in the conditional model above. As you can see in Table 2.7, the within-study coefficient for overall conceptual encoding accuracy was a significant predictor of accuracy solving right-blank problems, \(\hat{\beta} = .16, p = .04\). However, encoding accuracy was not a significant predictor of accuracy solving left-blank problems, \(\hat{\beta} = -.05, p = .53\). These results are in line with the run-on schema account.

2.3.5 Incorrect Strategy Analyses

If left-blank problems are more difficult to solve correctly because children interpret them as run-on traditional addition problems, then the add-to-equal-sign strategy should be the most common incorrect strategy used on left-blank problems. To examine this possibility, I tabulated each solution code for all incorrect problems solved by any children in the full sample. Because 175 children had perfect performance, this analysis included data from 1,239 children. Across all studies, 2,381 left-blank problems were solved incorrectly and 2,219 right-blank problems were solved incorrectly. The percentage of add-to-equal-sign strategies out of all incorrect strategies was 61% on left-blank problems compared to only 16% on right-blank problems. These results provide additional support for the run-on schema account.
2.3.6 Arithmetic Error Analyses

If the difficult unknown account provided an explanation for this result, then children should be more likely to make an arithmetic error on left-blank problems than on right-blank problems. I operationalized an arithmetic error to be 1 or 2 away from the correct solution (e.g., if the correct solution is 14, putting either 12, 13, 15, or 16 in the blank) when a child explains a correct strategy. Although the number of problems explained correctly was infrequent in the overall sample, I recorded each instance of an arithmetic error on each problem type in order to test whether a difference existed between problem formats. For left-blank problems, 14 out of 176 correct explanations (8%) included an arithmetic error for the solution, and for right-blank problems, 21 out of 231 correct explanations (9%) included an arithmetic error for the solution. This difference was not statistically significant, \( p = .68 \), indicating no clear support for the difficult unknown account.

2.4 Discussion

This study provides evidence that the role of encoding in children’s formal understanding of math equivalence depends on problem format. Specifically, contrary to the prevailing view that encoding performance predicts solving performance, accuracy on right-blank problems was significantly higher than accuracy on left-blank problems. Furthermore, overall encoding accuracy was significantly correlated with performance on right-blank problems, but not left-blank problems. Results suggest that the interpretation of what features are encoded may matter more than accuracy encoding problem features, and findings point to a common misconception of a run-on schema when children approach left-blank problems.
Previous research has found gender to be a significant predictor of children’s performance solving math equivalence problems, in that girls were less likely to solve problems correctly than boys (Hornburg, Rieber, & McNeil, 2017). Therefore, gender was included in the conditional models in this study to be able to see the effect of encoding accuracy when controlling for gender. Though this study was not designed to examine differential effects of gender, results revealed the impact of problem format on the role of gender in children’s performance. Specifically, gender was a significant predictor of the standardized mean difference in solving performance. In more interpretable terms, in the full sample, gender was a significant predictor of accuracy solving right-blank problems, but it was not a significant predictor of accuracy solving left-blank problems. It may be that girls are at even more of a disadvantage for right-blank problems, in which boys are taking the risk to try out a new strategy for a problem that cannot be conceptualized as a traditional math problem, but girls are still avoiding trying new strategies. More research is needed to examine why this difference exists.

2.4.1 Theoretical Implications

These results advance theory about the role of encoding in children’s problem solving performance. For math equivalence problems, the common “correct encoding leads to correct solving” mantra does not apply. In fact, I found evidence pointing to children’s incorrect run-on schema as the cause of incorrect responses on particular problem types — left-blank problems — that are generally encoded quite accurately. In fact, McNeil and Alibali (2004) demonstrated that children are just as accurate at encoding left-blank problems as they are at encoding traditional format problems. This supports the notion that feature knowledge may be more important for problem-solving
success than feature encoding, as demonstrated by Booth and Davenport (2013). In fact, children can encode left-blank problems entirely accurately, but because the format with “=” and “_ _” placed together is compatible with their existing knowledge “a + b = __,” then “a + b = __ + c” just gets assimilated into children’s established framework of computing left to right (Kuhn, 1989).

Results also add to the change-resistance account of children’s difficulties with math equivalence (McNeil & Alibali, 2005b). Previous hypotheses stemming from the change-resistance account focused on the positive association between encoding and solving, and predicted that because right-blank problems activate children’s narrow traditional arithmetic knowledge and are inaccurately encoded (McNeil & Alibali, 2004), and thus, be more difficult to solve correctly (McNeil & Alibali, 2004). Instead, my results support a hypothesis from the change-resistance account that is independent from the positive encoding-solving link assumption, and point to an additional way that children’s traditional arithmetic knowledge can be activated, in their misconceptualization of left-blank problems as run-on traditional addition problems. For both math equivalence problem formats, children are exhibiting change-resistance and adherence to the traditional arithmetic problem format. McNeil and Alibali (2005b) theorized that children’s everyday experiences with traditional arithmetic (e.g., a + b + c =__) activate three operational patterns (the “perform all operations on all given numbers” problem-solving strategy, the “operations = answer” problem structure, and the interpretation that the equal sign means “the total”), and found that knowledge of these operational patterns hindered children’s ability to learn from instruction. My results suggest there could be a fourth operational pattern picked up from arithmetic instruction:
the conceptualization of a run-on schema in which the equal sign can signal both “put the answer” and “begin the next problem.”

2.4.2 Practical Implications

These findings are of importance to both researchers and educators alike, as understanding of math equivalence is widely regarded as one of the most important concepts for developing children’s algebraic thinking (Blanton & Kaput, 2005; Carpenter et al., 2003; NCTM, 2000), and algebra literacy is critically important for students’ future academic and life success (Moses & Cobb, 2001; NRC, 1998). This difference in performance by format as a result of poor interpretation of what is encoded (rather than encoding inaccuracy) underscores the importance of introducing pre-algebra concepts early in the elementary grades. Indeed, many researchers and policymakers have called for the “algebrification” of early arithmetic, in which students discuss and reflect upon generalizations about properties of arithmetic (e.g., Blanton & Kaput, 2005; Carpenter et al., 2003; National Mathematics Advisory Panel, 2008; Schliemann, Carraher, & Brizuela, 2007; Stephens et al., 2015). Efforts to infuse conceptually-focused instruction in the classroom through professional development for teachers have been successful to promote early algebraic thinking (e.g., Jacobs, Franke, Carpenter, Levi, & Battey, 2007; Schliemann et al., 2007; Stephens et al., 2015). This is especially important given that many teachers are not necessarily aware of children’s entrenched misconceptions about the equal sign (Asquith, Stephens, Knuth, & Alibali, 2007; Stephens, 2006). Though an understanding of math equivalence has been included in the first and second grade Common Core State Standards (NGA Center & CCSSO, 2010), it is not clear that teachers are adapting classroom practices with respect to arithmetic practice to support
students’ understanding of math equivalence (Davenport, Kao, & McNeil, 2016; Silla, 2017).

Furthermore, with the evidence pointing towards the run-on schema account as the source of children’s difficulties solving left-blank problems, educators should be careful to not use the equal sign in ways that enforce that run-on schema (Carpenter et al., 2003). Unfortunately, second grade arithmetic lessons in the U.S. today can enforce that mindset of seeing the equal sign like a “perform the calculation” symbol. As recently as 2017, a classroom was observed in which the teacher wrote out a string of calculations using multiple equal signs as a way to demonstrate connections between multiplication facts (e.g., to help the class figure out $9 \times 6 = \_\_\_$, she wrote $10 \times 6 = 60$ and then continued the problem to complete a run-on equation $10 \times 6 = 60 - 6 = 54$) (Silla, 2017). Even in middle school, students are hindered by their left-to-right mindset. Though they may be able to solve math equivalence problems correctly and accept problems with operations on both sides as valid equations, seventh and eighth grade students still struggle with understanding the order of operations and fail to see the mistake with writing $2 + 1 \times 5 = 3 \times 4 + 3$, because they are so focused on performing left-to-right operations (Herscovics & Kieran, 1980). It may be possible that a focus on preventing the entrenchment of the run-on schema may aid in future understanding of order of operations as well.

2.4.3 Limitations and Future Directions

The present study is not without limitations. First, although great care was taken to accumulate a large amount of data for analysis, I still only had 14 total studies, ten of which included an encoding assessment. In the conditional models above, there were only
four degrees of freedom for the between-study effects, which limited the power to detect an existing effect. Furthermore, for some fine-grained analyses including those children who solved at least one problem correctly but not all correctly (and also had data from an encoding assessment), the subsample was fairly small ($n = 133$). In addition, because I limited the sample to problems solved prior to any traditional problems and problems encoded prior to any traditional problems, data for some children included only one problem of each type.

Second, although follow-up analyses point to children’s run-on schema as the driving force behind the difference in performance observed, there could be factors that affect problem-solving performance on each problem type that were not tested in the present study. For example, particular curricula may support problem solving of one type over another. The Everyday Mathematics curriculum developed by the University of Chicago School Mathematics Project and published by McGraw-Hill Education may be one that impacts performance, because it includes more instances of nontraditional problem formats (e.g., $a = b + _-$) than other curricula. The setting in which problems are solved likely plays a role as well, and I found some evidence for that, as within the subsample of children who solved at least one problem correctly but not all correctly, the assessment method (paper or pencil vs. whiteboard) was a significant predictor of performance on right-blank problems, but not on left-blank problems. This impact of problem-solving setting could be examined further in future studies. Other factors to be addressed in future research are children’s level of confidence or uncertainty when solving each problem type as well as the time it takes to solve the problems, which I was unable to examine in the present study.
Although these results demonstrate that right-blank problems are easier for children to solve correctly, past research shows that left-blank problems are encoded more accurately (McNeil & Alibali, 2004). It remains unclear which of the two types will be best for teaching children how to solve math equivalence problems correctly. Therefore, I tested this in Study 2.
2.5 Figures

Figure 2.1: Accuracy solving left-blank (top graph) and right-blank problems (bottom graph) for the full sample ($N = 1,414$).
Figure 2.2: Accuracy solving left-blank (top graph) and right-blank problems (bottom graph) for the subsample ($n = 196$).
Figure 2.3: Overall encoding accuracy for the full sample (top graph; $N = 981$) and the subsample (bottom graph; $n = 133$).
### TABLE 2.1

INTEGRATIVE DATA ANALYSIS: STUDY CHARACTERISTICS

<table>
<thead>
<tr>
<th>Study</th>
<th>Authors</th>
<th>Data Collection</th>
<th>Sample Size (N)</th>
<th>Method</th>
<th>Grade M (SD)</th>
<th>Percent Girls</th>
<th>Percent in Lab</th>
<th>Percent in Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alibali et al. (2017)</td>
<td>1997-1999</td>
<td>113</td>
<td>Board</td>
<td>4.00 (.00)</td>
<td>50</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>Hattikudur &amp; Alibali (2010)</td>
<td>2005-2007</td>
<td>111</td>
<td>Board</td>
<td>3.23 (.43)</td>
<td>50</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>McNeil et al. (2011)</td>
<td>2008</td>
<td>30</td>
<td>Board</td>
<td>2.07 (.25)</td>
<td>50</td>
<td>23</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>McNeil et al. (2012)</td>
<td>2008-2009</td>
<td>35</td>
<td>Board</td>
<td>2.74 (.44)</td>
<td>54</td>
<td>63</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>Chesney et al. (2012)</td>
<td>2009</td>
<td>37&lt;sup&gt;be&lt;/sup&gt;</td>
<td>Board</td>
<td>2.39 (.49)</td>
<td>59</td>
<td>60</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>Cook, Duffy, &amp; Fenn (2013)</td>
<td>2011</td>
<td>232</td>
<td>Paper</td>
<td>2.56 (.55)</td>
<td>52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>McNeil et al. (2017)</td>
<td>2011-2013</td>
<td>135</td>
<td>Board</td>
<td>2.02 (.15)</td>
<td>41</td>
<td>41</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>Fyfe &amp; Rittle-Johnson (2016), Exp. 1</td>
<td>2012-2013</td>
<td>159</td>
<td>Paper</td>
<td>2.60 (.49)</td>
<td>59</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>9</td>
<td>Byrd et al. (2014)</td>
<td>2013</td>
<td>99</td>
<td>Board</td>
<td>2.98 (.14)</td>
<td>52</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Hall (2014)</td>
<td>2013</td>
<td>50</td>
<td>Board</td>
<td>2.02 (.14)</td>
<td>52</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Byrd et al. (2015)</td>
<td>2013</td>
<td>164</td>
<td>Paper</td>
<td>2.00 (.00)</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>Hornburg et al. (2015)</td>
<td>2014-2015</td>
<td>107</td>
<td>Board</td>
<td>2.00 (.24)</td>
<td>37</td>
<td>51</td>
<td>95</td>
</tr>
</tbody>
</table>

<sup>a</sup>Participants from control condition only.
<sup>b</sup>Grade is missing for one child in this study.
<sup>c</sup>One child in this study has missing data for problem encoding.
<sup>d</sup>Two children in this study have missing data for gender.
## TABLE 2.2

INTEGRATIVE DATA ANALYSIS: PROBLEMS GIVEN IN EACH STUDY

<table>
<thead>
<tr>
<th>Study or Studies</th>
<th>Problem Solving Equations in Order</th>
<th>Problem Encoding Equations in Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 and 13</td>
<td>[4 + 3 + 6 = _ + 6]</td>
<td>6 + 8 + 9 = _ + 9</td>
</tr>
<tr>
<td></td>
<td>[3 + 9 + 5 = _ + _]</td>
<td>5 + 6 + 9 = _ + _</td>
</tr>
<tr>
<td>2</td>
<td>[7 + 5 + 3 = 7 + _]</td>
<td>9 + 6 + 8 = 9 + _</td>
</tr>
<tr>
<td></td>
<td>[3 + 8 + 6 = _ + 6]</td>
<td>6 + 3 + 8 = 6 + _</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 + 3 + 7 = 4 + _</td>
</tr>
<tr>
<td>3, 4, and 5</td>
<td>[2 + 7 = 6 + _]</td>
<td>[7 + 1 = _ + 6]</td>
</tr>
<tr>
<td>(random set)</td>
<td>[1 + 5 = _ + 2]</td>
<td>[4 + 5 = 3 + _]</td>
</tr>
<tr>
<td></td>
<td>[3 + 5 + 6 = 3 + _]</td>
<td>[7 + 1 = _ + 6]</td>
</tr>
<tr>
<td></td>
<td>[7 + 2 + 4 = _ + 4]</td>
<td>[7 + 1 = _ + 6]</td>
</tr>
<tr>
<td></td>
<td>[3 + 5 + 6 = 3 + _]</td>
<td>[4 + 5 = 3 + _]</td>
</tr>
<tr>
<td></td>
<td>[5 + 9 + 8 = 5 + _]</td>
<td>[4 + 5 = 3 + _]</td>
</tr>
<tr>
<td>6</td>
<td>[1 + 5 = _ + 2]</td>
<td>[7 + 1 = _ + 6]</td>
</tr>
<tr>
<td></td>
<td>[2 + 7 = 6 + _]</td>
<td>[2 + 3 + 6 = 2 + _]</td>
</tr>
<tr>
<td></td>
<td>[7 + 2 + 4 = _ + 4]</td>
<td>[3 + 5 + 4 = _ + 4]</td>
</tr>
<tr>
<td></td>
<td>[3 + 5 + 6 = 3 + _]</td>
<td>[4 + 5 = 3 + _]</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>7, 10, and 12</td>
<td>[1 + 5 = _ + 2]</td>
<td>[4 + 5 = 3 + _]</td>
</tr>
<tr>
<td>(Only 7 and</td>
<td>[2 + 7 = 6 + _]</td>
<td>[7 + 1 = _ + 6]</td>
</tr>
<tr>
<td>12 had encoding)</td>
<td>[7 + 2 + 4 = _ + 4]</td>
<td>[2 + 3 + 6 = 2 + _]</td>
</tr>
<tr>
<td></td>
<td>[3 + 5 + 6 = 3 + _]</td>
<td>[3 + 5 + 4 = _ + 4]</td>
</tr>
<tr>
<td>Study or Studies</td>
<td>Problem Solving Equations in Order</td>
<td>Problem Encoding Equations in Order</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>8(^a)</td>
<td>1 + 5 = __ + 2&lt;br&gt;7 + 2 + 4 = __ + 4&lt;br&gt;2 + 7 = 6 + __&lt;br&gt;3 + 5 + 6 = 3 + 9</td>
<td>7 + 1 = __ + 6&lt;br&gt;3 + 5 + 4 = __ + 4&lt;br&gt;4 + 5 = 3 + __&lt;br&gt;2 + 3 + 6 = 2 + 3</td>
</tr>
<tr>
<td>9 and 14</td>
<td>5 + 4 + 6 = __ + 6&lt;br&gt;3 + 5 + 9 = __ + 9&lt;br&gt;8 + 4 + 3 = 8 + __&lt;br&gt;7 + 5 + 8 = 7 + __&lt;</td>
<td>N/A</td>
</tr>
<tr>
<td>11</td>
<td>5 + 4 = __ + 4&lt;br&gt;8 + 2 = __ + 6&lt;br&gt;7 + 2 + 4 = __ + 4&lt;br&gt;7 + 4 + 6 = __ + 3&lt;br&gt;2 + 6 = 2 + __&lt;br&gt;1 + 5 = 2 + __&lt;br&gt;3 + 5 + 6 = 3 + __&lt;br&gt;6 + 2 + 8 = 4 + __&lt;</td>
<td>4 + 5 = 3 + __&lt;br&gt;7 + 1 = __ + 6&lt;br&gt;2 + 3 + 6 = 2 + __&lt;br&gt;3 + 5 + 4 = __ + 4</td>
</tr>
</tbody>
</table>

\(^a\)Empty boxes were used instead of blanks in this study.
### TABLE 2.3

**EXAMPLE PROBLEM-SOLVING STRATEGIES AND EXPLANATIONS**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Solution</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add all</td>
<td>25</td>
<td>“I added 5, 4, 7, and 7.”</td>
</tr>
<tr>
<td>Add to equal sign</td>
<td>16</td>
<td>“I added 5 plus 4 plus 7 to get 16.”</td>
</tr>
<tr>
<td>Add to equal sign &amp; add all</td>
<td>16 &amp; 25</td>
<td>“I added 5 plus 4 plus 7 and it was 16 and then I added the 7 so I got 25.”</td>
</tr>
<tr>
<td>Carry</td>
<td>5</td>
<td>“There was a 5 here, so I put a 5 over here too.”</td>
</tr>
<tr>
<td>Add two numbers</td>
<td>11</td>
<td>“I added 4 and 7.”</td>
</tr>
<tr>
<td>Make right side equal to the number in front of the equal sign</td>
<td>0</td>
<td>“I put 0 because 7 plus 0 equals 7.”</td>
</tr>
<tr>
<td>Idiosyncratic</td>
<td>2</td>
<td>“Because 5 plus 2 equals 7.”</td>
</tr>
<tr>
<td>Equalize</td>
<td>9</td>
<td>“5 plus 4 plus 7 equals 16 and 9 plus 7 equals 16.”</td>
</tr>
<tr>
<td>Add-subtract</td>
<td>9</td>
<td>“I added 5 plus 4 plus 7 and then took away 7 to get 9.”</td>
</tr>
<tr>
<td>Grouping</td>
<td>9</td>
<td>“I saw 7s on both sides, so I just added 5 plus 4 and put 9.”</td>
</tr>
<tr>
<td>Study No.</td>
<td>No. of Problems</td>
<td>Percent RB First</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. RB = Right-blank problems; LB = Left-blank problems.
TABLE 2.5

INTEGRATIVE DATA ANALYSIS: PROBLEM ENCODING PERFORMANCE BY STUDY

<table>
<thead>
<tr>
<th>Study No.</th>
<th>No. of Problems</th>
<th>Mean Conceptual Acc. (SD) Overall</th>
<th>Mean Conceptual Acc. (SD) on RB</th>
<th>Mean Conceptual Acc. (SD) on LB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>.54 (.40)</td>
<td>.46 (.50)</td>
<td>.61 (.49)</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>.57 (.40)</td>
<td>.57 (.40)</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>.29 (.35)</td>
<td>.28 (.36)</td>
<td>.30 (.39)</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>.30 (.38)</td>
<td>.23 (.37)</td>
<td>.37 (.43)</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>.45 (.40)</td>
<td>.36 (.44)</td>
<td>.54 (.44)</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>.41 (.38)</td>
<td>.36 (.40)</td>
<td>.47 (.45)</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>.60 (.37)</td>
<td>.52 (.42)</td>
<td>.68 (.42)</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>.30 (.34)</td>
<td>.25 (.35)</td>
<td>.35 (.41)</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>.44 (.41)</td>
<td>.37 (.43)</td>
<td>.52 (.46)</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>.40 (.41)</td>
<td>.38 (.49)</td>
<td>.43 (.50)</td>
</tr>
</tbody>
</table>

*Note.* RB = Right-blank problems; LB = Left-blank problems.
### TABLE 2.6

**INTEGRATIVE DATA ANALYSIS: PREDICTING COHEN’S D**

<table>
<thead>
<tr>
<th></th>
<th>Between-Study ($df = 4$)</th>
<th>Within-Study ($df = 967$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method; 0 = Paper, 1 = Board</td>
<td>$b = -0.01$</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>$SE = 0.10$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = .89$</td>
<td></td>
</tr>
<tr>
<td>Number of Encoding Problems</td>
<td>$b = -0.17$</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>$SE = 0.10$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = .15$</td>
<td></td>
</tr>
<tr>
<td>Grade; 0 = Boy, 1 = Girl</td>
<td>$b = -0.09$</td>
<td>$b = -0.001$</td>
</tr>
<tr>
<td></td>
<td>$SE = 0.14$</td>
<td>$SE = 0.07$</td>
</tr>
<tr>
<td></td>
<td>$p = .55$</td>
<td>$p = .99$</td>
</tr>
<tr>
<td>Gender; 0 = Boy, 1 = Girl</td>
<td>$b = -0.15$</td>
<td>$b = -0.10$</td>
</tr>
<tr>
<td></td>
<td>$SE = 0.61$</td>
<td>$SE = 0.04$</td>
</tr>
<tr>
<td></td>
<td>$p = .81$</td>
<td>$p = .02$</td>
</tr>
<tr>
<td>Overall Conceptual Encoding</td>
<td>$b = -0.44$</td>
<td>$b = 0.14$</td>
</tr>
<tr>
<td>Accuracy</td>
<td>$SE = 0.36$</td>
<td>$SE = 0.07$</td>
</tr>
<tr>
<td></td>
<td>$p = .30$</td>
<td>$p = .04$</td>
</tr>
</tbody>
</table>

*Note.* Predictors significant at $p < .05$ are bolded.
<table>
<thead>
<tr>
<th></th>
<th>Predicting RB Accuracy</th>
<th></th>
<th>Predicting LB Accuracy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Between-Study</td>
<td>Within-Study</td>
<td>Between-Study</td>
<td>Within-Study</td>
</tr>
<tr>
<td></td>
<td>($df = 4$)</td>
<td>($df = 120$)</td>
<td>($df = 4$)</td>
<td>($df = 120$)</td>
</tr>
<tr>
<td><strong>Method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = Paper, 1 = Board</td>
<td>$b = 0.28$</td>
<td>$SE = 0.08$</td>
<td>N/A</td>
<td>$b = 0.05$</td>
</tr>
<tr>
<td><strong>$p = .02$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of Encoding Problems</strong></td>
<td>$b = -0.09$</td>
<td>$SE = 0.12$</td>
<td>N/A</td>
<td>$b = 0.34$</td>
</tr>
<tr>
<td><strong>$p = .49$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td>$b = -0.03$</td>
<td>$SE = 0.16$</td>
<td>$b = 0.25$</td>
<td>$b = 0.05$</td>
</tr>
<tr>
<td><strong>$p = .86$</strong></td>
<td>$b = -0.12$</td>
<td>$SE = 0.12$</td>
<td>$SE = 0.17$</td>
<td>$SE = 0.10$</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = Boy, 1 = Girl</td>
<td>$b = 0.17$</td>
<td>$SE = 0.30$</td>
<td>$b = -0.10$</td>
<td>$b = 0.12$</td>
</tr>
<tr>
<td><strong>$p = .61$</strong></td>
<td>$b = -0.17$</td>
<td>$SE = 0.08$</td>
<td>$SE = 0.33$</td>
<td>$SE = 0.06$</td>
</tr>
<tr>
<td><strong>Overall Conceptual Encoding Accuracy</strong></td>
<td>$b = 0.67$</td>
<td>$SE = 0.28$</td>
<td>$b = 0.45$</td>
<td>$b = -0.05$</td>
</tr>
<tr>
<td><strong>$p = .07$</strong></td>
<td>$b = 0.16$</td>
<td>$SE = 0.08$</td>
<td>$SE = 0.31$</td>
<td>$SE = 0.08$</td>
</tr>
</tbody>
</table>

*Note.* Predictors significant at $p < .05$ are bolded.
CHAPTER 3:

STUDY 2 – EXAMINING THE ROLE OF ENCODING IN CHILDREN’S LEARNING

3.1 Method

3.1.1 Participants

Participants were third graders from local schools, after-school programs, and summer enrichment programs in the Midwestern United States, recruited through letters and permission slips sent to parents. I chose to work with third graders because by the third grade, children are likely to have robust misconceptions about the meaning of the equal sign (McNeil, 2007). I recruited third graders in the Summer and Fall of 2016. Students who participated in the summer had either just finished third grade or were entering third grade in the Fall. The main analysis had three groups (two experimental and one control), so my goal was to obtain a total sample size of 96 children, to detect a medium-to-large effect size ($f = .325$) at an alpha of 0.05 and a desired power of 0.80 (Faul, Erdfelder, Buchner, & Lang, 2013). For the purposes of the main analysis, children who solved any of the math equivalence problems on the pretest correctly were excluded (because I was interested in the effects of the interventions on learning), thus, I recruited more participants with the expectation that about 20-25% of children would be able to solve at least one math equivalence problem correctly. In fact, I continued to recruit more participants than I originally planned in order to hit the benchmark number who were not able to solve any problems correctly at the pretest. In total, I worked with 181 children.
and 76 (42%) of them solved at least one math equivalence problem correctly at a pretest, a much higher percentage than expected.

The final sample included 105 children (40 boys, 65 girls; $M$ age = 8 years, 10 months, $SD = 6$ months, min = 7 years 11 months, max = 11 years 6 months). Most children (97%) were either 8 or 9 years old; one child was not yet 8 years old and two children were already at least 10 years old (conclusions remain unchanged when these children are excluded). The race/ethnicity of the sample was 50% white, 28% Hispanic or Latino, 20% African American or black, 1% Asian, and 1% other.

Sessions were conducted in a quiet room at a research lab or local school or community center. All recruiting and research procedures were consistent with APA ethical guidelines. In addition to parental consent, verbal assent was obtained from the child before each session. Children were also free to choose to stop participating at any time, and experimenters monitored children’s non-verbal behavior to evaluate whether they were upset and may have wanted to end the session. Each child received some small gifts (e.g., colorful pencils, erasers, stickers) in exchange for their participation, and parents (or schools/programs) were compensated for their time and allowance of child participation.

3.1.2 Measures

3.1.2.1 Pretest

Children completed a paper-and-pencil pretest consisting of a variety of math problems, including problems to solve and problems to reconstruct after viewing them briefly (i.e., encoding problems, described in the following section). Instructions for each
problem were read aloud to the child. Problems included traditional arithmetic problems, math equivalence problems, geometry items, and probability items. General math achievement items were chosen or adapted from the National Assessment of Educational Progress (NAEP) fourth grade items determined easiest for children, to approximate a third-grade level. The four math equivalence problems were embedded within a larger mathematics assessment in order to reduce pretest sensitization to math equivalence problems. Children were randomly assigned to receive a pretest in which the first math equivalence problem seen is a left-blank problem (set A) or a right-blank problem (set B). The exact problems presented can be found in Table 3.1. For coding, correctness was determined by the number children wrote in the blank. As in prior work in which children were not asked to explain their solutions to math equivalence problems, solutions that were within ±1 of the correct solution were coded as correct only if the child had solved another problem exactly correct. Because children were not asked to explain solutions (in an effort to avoid focused attention on math equivalence problems prior to the intervention), care was taken in the process of selecting problems for the pretest so that no solutions that were within ±1 of the correct solution overlapped with any other traditional incorrect strategy code.

During the encoding section (set in the middle of the pretest), children were asked to reconstruct four problems with operations on both sides of the equal sign (e.g., 7 + 4 + 5 = 7 + ___) using paper and pencil after viewing each for five seconds (cf. Chase & Simon, 1973; Siegler, 1976). Intermixed with these math equivalence problems was a traditional addition problem for children to encode. Again, the order of problems (left-blank first vs. right-blank first) depended on whether the child was assigned to set A or
set B. The exact problems presented can be found in Table 3.1. The experimenter told children that they did not have to solve the problems; rather, they just needed to write exactly what they saw after the experimenter put the problem down. Children’s responses were coded based on a system from previous research, and was the same system used in Study 1 (e.g., McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999). Using this coding system, children’s errors can be categorized as number errors, such as mixing up the order of the numbers, and as conceptual errors, such as omitting the equal sign, omitting the plus on the right side of the equal sign, or converting the problem to a traditional arithmetic problem.

3.1.2.2 Task Difficulty Rating

To assess children’s level of cognitive load while solving the problems, I included a brief measure of task difficulty, adapted from the first item of the cognitive load assessment used by Fyfe, DeCaro, and Rittle-Johnson (2015) (based on Paas, 1992; see also Paas, Tuovinen, Tabbers, and Van Gerven, 2003 for evidence of similar subjective rating scales used with adolescents and adults). I decided to use only a measure of task difficulty as an index of cognitive load and not measures of mental effort or mental frustration because task difficulty scales in past research with children have been found to be the most reliable and valid (Fyfe & Rittle-Johnson, 2016) After solving the last problem of the intervention, children were given a sheet of paper with a question and a 5-point rating scale and the experimenter read the question, “How easy or difficult was it to solve all of those problems?” The experimenter read each of the response options aloud (ranging from very easy to very difficult) and asked children to circle whichever one best matched what they thought. Children in the control condition were given a similar task,
with the singular question, “How easy or difficult was it to solve that problem?” following the one math equivalence problem they solved at the whiteboard. This type of rating measure has been found to be both predictive of learning outcomes and sensitive to changes in cognitive load in work with undergraduates (Wiebe, Roberts, & Behrend, 2010). The exact task for both the experimental conditions and control condition is displayed in Figure 3.1.

3.1.2.3 Posttest, Opposite Format, and Other Format Problems

After the intervention, children completed an immediate posttest. First, children completed a problem encoding task. The problem encoding protocol was the same as the pretest, with a fixed order of math equivalence problems intermixed with a traditional problem, given in order either set A or set B. The exact problems presented can be found in Table 3.2. Coding of errors on these problems was identical to the coding scheme used for the pretest.

Next, children solved six math equivalence problems on paper, three of the format in which they were instructed (i.e., posttest problems) followed by three problems of the opposite format. Children in the control condition were first presented with three problems of the format they solved at the end of the intervention phase. Each problem was presented on a separate page. The exact problems presented can be found in Table 3.2. For this task, the experimenter said, “Try your best to solve the problem, and then write the number that goes in the blank.” After children solved all six problems in the packet, the experimenter went back to the first problem and asked, “Can you tell me how you got x?” (x denotes the given solution) (cf. Alibali, 1999; Perry, 1991; Rittle-Johnson & Alibali, 1999; Siegler, 2002), and continued asking for explanations for each
subsequent problem. This method was used rather than asking children to explain directly after solving each problem so that there would be no effect of self-explanation on subsequent performance. Strategies were coded using a system from previous research (e.g., McNeil & Alibali, 2004; Perry et al., 1988). The strategy was first coded by what the solution alone indicated. Responses were coded as reflecting a particular strategy as long as they were within ±1 of the response that would be achieved with that particular strategy. If the solution in the blank was unclear, then the child’s verbal strategy was considered. Table 2.3 presents common strategies used to solve math equivalence problems with example verbal explanations. Children were free to change what number they wrote in the blank, but correctness was determined by the final number they put in the blank before going through and explaining their solutions, in order to avoid the potential impact of seeing the other problem type and/or self-explanation on performance. Because these problems included four different addends and some high sums to calculate for each side of the equation (up to 20), great care was taken to examine correct verbal explanations which accompanied responses that were more than one away from the correct solution. For analysis, if the response given was not within ±1 of a traditional incorrect strategy (add-all or add-to-equal-sign), arithmetic errors greater than one were allowed (i.e., the response was counted as correct) when accompanied by a clearly correct verbal strategy. Note, however, that conclusions are unchanged when a very strict approach is taken instead.

Finally, children solved three problems on paper that differed from the other posttest and opposite format problems in terms of blank position (e.g., blank in the middle position between two addends, blank on the left side of the equal sign). As with
the pretest, the random assignment to receive set A or set B determined whether children saw a problem of the format \( _- + a = b + c + d \) vs. \( a + _- = b + c + d \) first after the middle-blank problem. The exact problems presented can be found in Table 3.2. After solving all three of these problems, children were asked to explain each of their solutions, in a similar procedure as described above.

3.1.3 Procedure

I used a pretest-intervention-posttest design. After the posttest, all children received a brief lesson on math equivalence tailored specifically to their needs. The session lasted approximately 40 minutes. All sessions were videotaped, so that I could study the strategies that children used when solving the problems. See Appendix A for the complete script used during the one-on-one session.

3.1.3.1 Experimental Conditions

Children were randomly assigned to one of three conditions for instruction: left-blank instruction \((n = 36)\), right-blank instruction \((n = 34)\), or a control condition \((n = 35)\). There were no differences among conditions in terms of age, gender, or session location, \(ps > .88\). The only aspect that varied between the two experimental conditions is the format in which the math equivalence problems were written during the instruction: either “left-blank” (i.e., \( a + b + c = _- + d \)) or “right-blank” (i.e., \( a + b + c = d + _- \)).

Children in the two experimental conditions were given instruction on how to solve math equivalence problems. During the instruction, children received a sequence of hints (based on previous intervention studies) that increased in the level of scaffolding to help them solve problems correctly. This scaffolding was designed with the natural
tutoring setting in mind, because the intent was to have this tutoring be applicable to the real world. Each level of scaffolding built on the previous one, so that by the end there was a cumulative impact of several hints on the last problem. First, children were given a problem to solve and given feedback after solving the problem. Children had the opportunity to try again on that problem if incorrect. For the second problem, the equal sign was presented in red ink. For the third problem, children were asked to point to the equal sign before solving. In the next level, children were asked what they thought the equal sign means (cf. Baroody & Ginsburg, 1983; Behr et al., 1980), were told the relational meaning of the equal sign, and were asked to repeat that relational definition prior to attempting to solve the third problem again. After the definition level, children were taught a correct strategy known as the equalize strategy. The goal of this strategy is to determine what number must go in the blank so that both sides have the same amount. It has been taught in several previous studies because it is thought to promote a deeper conceptual understanding beyond simply that procedure (e.g., Broaders et al., 2007; Cook et al., 2013; Cook, Mitchell, & Goldin-Meadow, 2008). The equalize strategy is also the correct strategy children tend to generate first on their own (e.g., Alibali et al., 2009). On the fourth problem of the intervention, children were first taught the equalizer strategy as a conceptual tool, without step-by-step assistance. If children were incorrect, the final level of scaffolding was to guide children through adding the specific numbers on the left side of the problem to help them along with implementing the strategy on that fourth problem. Finally, the fifth problem presented was similar to the first problem, with black ink and no additional space around the equal sign, as an index of understanding how to solve the problems correctly in the same context as instruction (i.e., at a whiteboard). The
exact problems presented during the intervention can be found in Table 3.3. See Figure 3.2 for screenshots of the tutor during each scaffolding level.

Children in the control condition were taught strategies for completing an unrelated task (a word search containing non-math words) during the intervention phase, to provide a comparison group that did not receive equalizer strategy instruction. As a parallel to the increased levels of help given to the experimental groups, children first were given two minutes to work on the word search on their own. Then the experimenter aided in the search process, provided encouragement, and described helpful strategies such as crossing off words that are found, looking for double letters in words, and using a circle search around the first letter of a word. After 10 minutes had passed, children were given one math equivalence problem to solve up at a whiteboard, with no assistance, followed by the task difficulty rating. Children in the control condition were randomly assigned to either solve a left-blank problem or a right-blank problem, the same problem as the last one solved in each experimental condition. All problem-solving strategies were coded from the same coding scheme as described previously. The total time (in minutes) elapsed in the intervention phase was comparable among the left-blank ($M = 11.63, SD = 4.57$), right-blank ($M = 10.76, SD = 3.51$), and control conditions ($M = 11.46, SD = 1.22$), $p = .85$.

3.2 Results

3.2.1 Pretest Performance

Recall that I was interested in the effects of the interventions on learning, so the main analyses are limited to the 105 children who were not able to solve any of the math
equivalence problems correctly at the pretest. Also recall that the general math achievement items were adapted from items in the NAEP assessment labeled as easiest for fourth graders, to approximate a third-grade level. However, they still were quite challenging for this sample, with children solving an average of 2.79 (out of 6, $SD = 1.55$) problems correctly. Proficiency on the general math items was significantly correlated with age, $r = .34$, $p < .001$. There were no differences among conditions in terms of performance on the general math achievement items, $p = .52$. Not surprisingly, children who solved at least one math equivalence problem correctly ($n = 75$, excluded from the main analyses) performed significantly better on the math achievement items ($M = 3.88$, $SD = 1.64$) than children who did not solve any math equivalence problems correctly, $p < .001$.

In terms of encoding at pretest, 68% of children encoded at least one math equivalence problem correctly (free of conceptual errors). The average number of math equivalence problems encoded correctly (out of 4) was 1.41 ($SD = 1.28$), which translates to about 35% accuracy overall. As expected, average encoding accuracy for left-blank problems ($44\%, SD = 0.42$) was higher than average encoding accuracy for right-blank problems ($26\%, SD = 0.35$), $p < .001$. There were no differences among conditions in terms of accuracy on the left-blank encoding problems, $p = .39$, or right-blank encoding problems, $p = .89$.

3.2.2 Accuracy During and After the Intervention by Test Phase and Condition

Figure 3.3 displays average problem-solving accuracy across test phases by condition. Figures 3.4 and 3.5 display the problem solving data in separate distributions by test phase for each experimental condition (control condition accuracy was at floor,
except two children in the control condition solved one of the three problems of other formats correctly). Internal consistency reliability was $\alpha = .80$ for the five intervention problems, $\alpha = .98$ for the three posttest problems, $\alpha = .97$ for the three opposite format problems, and $\alpha = .79$ for the three problems of other formats.

I conducted a mixed-factor ANCOVA with condition as the between-subjects factor, test phase (intervention, posttest, opposite format, other formats) as the within-subjects factor, and accuracy as the dependent measure. I included pretest math proficiency as a covariate because it was a significant predictor of accuracy (and age was not), and power is improved when such prognostic covariates are used (Kahan, Jairath, Doré, & Morris, 2014).

This analysis yielded main effects of pretest math proficiency, $F(1, 101) = 24.67, p < .001, \eta_p^2 = .20$, condition, $F(2, 101) = 49.35, p < .001, \eta_p^2 = .49$, and test, $F(3, 303) = 6.52, p < .001, \eta_p^2 = .06$. These effects were qualified by the interactions between condition and test, $F(6, 303) = 4.64, p < .001, \eta_p^2 = .08$, and between pretest math proficiency and test, $F(3, 303) = 2.95, p = .03, \eta_p^2 = .03$. Accuracy in the right-blank (50%) and left-blank (50%) conditions were significantly higher than that in the control condition (0%) across all test phases, $p < .001$. In follow-up contrasts, I first compared the two experimental conditions to the control condition for each test phase. Average accuracy in the experimental conditions was significantly higher than in the control condition for the intervention (46% vs. 0%; $p < .001$), posttest (65% vs. 0%, $p < .001$), opposite format problems (50% vs. 0%; $p < .001$), and other format problems (38% vs. 1%; $p < .001$). Finally, I examined the contrast between the two experimental conditions for all test phases. Accuracy in the right-blank condition was significantly higher than in
the left-blank condition for the intervention phase (52% vs. 40%, \( p = .013 \)) but not for posttest (63% vs. 67%, \( p = .57 \)), opposite format problems (50% vs. 50%; \( p = .98 \)), or other format problems (35% vs. 41%; \( p < .38 \)). Note that when logistic regression is used instead of ANCOVA, with separate models predicting at least one problem correct and all problems correct for each test phase (controlling for pretest math proficiency), results are unchanged. Additionally, when the full sample of children who participated (\( N = 181 \)) is analyzed and pretest accuracy on math equivalence problems is included as an additional covariate, results are unchanged.

Figure 3.6 displays average posttest encoding accuracy for each problem format by condition. Figure 3.7 displays the encoding data in separate distributions by problem format and condition. In a similar analysis, I conducted a mixed-factor ANCOVA with condition as the between-subjects factor, problem format (intervention format, opposite format)\(^2\) as the within-subjects factor, posttest encoding accuracy as the dependent measure, and both pretest math proficiency and pretest overall encoding accuracy as covariates. There were significant main effects of pretest math proficiency, \( F(1, 100) = 5.73, p = .02, \eta_p^2 = .05 \), and pretest encoding accuracy \( F(1, 100) = 26.96, p < .001, \eta_p^2 = .21 \); however, there were not significant main effects of condition, \( F(2, 100) = 1.20, p = .31, \eta_p^2 = .02 \), or problem format, \( F(1, 100) = 0.32, p = .57, \eta_p^2 = .003 \). Interactions were not significant between pretest math proficiency and problem format, \( F(1, 100) = 1.27, p = .26, \eta_p^2 = .01 \), or between pretest encoding accuracy and problem format, \( F(1, 100) = 0.84, p = .36, \eta_p^2 = .01 \).

\(^2\) Recall that children in the control condition solved one math equivalence problem at the board, and were randomly assigned to receive either a left-blank or right-blank format problem.
Although the main effects of condition and problem format were not significant, the interaction between condition and problem format was significant, $F(2, 100) = 7.98, p = .001, \eta_p^2 = .14$. In follow-up contrasts, I first compared the two experimental conditions to the control condition for each problem format. Average posttest encoding accuracy in the experimental conditions was not significantly different from in the control condition for the type of problem seen during the intervention (49% vs. 56%; $p = .42$) or the opposite format problems (50% vs. 52%; $p = .82$). Finally, I examined the contrast between the two experimental conditions for each problem format. Average posttest encoding accuracy in the left-blank condition was significantly higher than in the right-blank condition for the problem format seen during the intervention (64% vs. 34%; $p = .001$) but not for the opposite problem format (45% vs. 55%, $p = .28$). Note that when the full sample of children who participated ($N = 181$) is analyzed and pretest accuracy on math equivalence problems is included as an additional covariate, results are unchanged.

3.2.3 Does Pretest Encoding Accuracy Moderate the Effect of the Interventions?

In order to determine whether encoding accuracy at pretest was a moderator of intervention effectiveness, I classified children in the experimental conditions ($n = 70$) as “accurate” or “inaccurate” encoders based on a median split. From this split, 28 children were classified as accurate encoders (50% or greater accuracy), and 42 children were classified as inaccurate encoders (less than 50% accuracy). Note that conclusions were unchanged when instead the median value (25% accuracy; $n = 18$) was included in the accurate group. In an ANCOVA with only the children in the experimental conditions, I added pretest encoder level (accurate or inaccurate) as a between-subjects variable in addition to condition, with test phase (intervention, posttest, opposite format, and other
formats) as the within-subjects factor, accuracy as the dependent measure, and pretest math proficiency as the covariate. This yielded main effects of pretest math proficiency, $F(1, 65) = 20.84, p < .001, \eta^2_p = .24$, and test, $F(3, 195) = 3.83, p = .01, \eta^2_p = .06$.

However, there was no significant main effect of condition, $F(1, 65) < .001, p = .99, \eta^2_p < .001$ or pretest encoder level, $F(1, 65) = 0.09, p = .76, \eta^2_p = .001$. The interactions between pretest math proficiency and test, $F(3, 195) = 2.07, p = .11, \eta^2_p = .03$, between pretest encoding level and test, $F(3, 195) = 2.09, p = .10, \eta^2_p = .03$, between condition and test, $F(3, 195) = 1.43, p = .24, \eta^2_p = .02$, or between pretest encoding level and condition, $F(1, 65) = 0.06, p = .81, \eta^2_p = .001$, each were not statistically significant. The three-way interaction between encoding level, condition, and test was also not significant, $F(3, 195) = 0.22, p = .88, \eta^2_p = .003$. Note that again when the full sample of children who participated ($N = 181$) is analyzed and pretest accuracy on math equivalence problems is included as an additional covariate, results are unchanged.

3.2.4 Levels of Scaffolding Needed to Learn

As a further examination into the interventions, I coded the percentage of children in each condition who learned to solve a problem correctly at each level of scaffolding. There were six levels of scaffolding: feedback, red equal sign, additional spacing and point to equal sign, equal sign definition, conceptual strategy, and explicit procedure.

Figure 3.8 displays a breakdown of level of scaffolding required by condition, presented as the percentage of children learning at each level, including those who did not learn from the intervention. As can be seen in the figure, the only children able to learn after the first two levels were in the right-blank condition, whereas children in the left-blank condition seemed to need the definition before a sizable number of them were able to
learn. Specifically, 27% of children in the right-blank condition were able to generate a correct strategy after receiving only minimal scaffolding (i.e., within the first two levels, feedback or red equal sign), whereas none of the children in left-blank condition did, $t = -3.45, p = .002$. Children’s level of scaffolding needed was coded as the level after which they first solved and explained a problem correctly; however, one child in the right-blank condition did not provide a correct explanation for the one problem he solved correctly so this child was given credit for the level at which she solved the problem correctly (conceptual strategy). Note that results are unchanged when this child was coded as never learning.\(^3\) As you can see in the figure, once more explicit conceptual support was provided, the children in left-blank condition seemed to catch up to their peers in the right-blank condition. Indeed, the percentage of children who were never able to solve a problem correctly during the intervention was not significantly different between the right-blank condition (12%) and the left-blank condition (19%), $t = 0.88, p = .39$.

Most children continued to solve problems correctly during the intervention after solving one correctly (92%), with the exceptions including the child mentioned above (he did not solve the procedure level problem or the final problem correctly) and those mentioned in the footnotes (although, after explaining the problem correctly at the procedure stage the child in the right-blank condition did maintain that correct strategy on the final problem of the intervention) and one child in each condition who solved one

\(^3\) One child in the left-blank condition generated a correct strategy on the final problem of the intervention (after the scaffolding sequence was complete and children were not given assistance), and was coded in the same category as those who solved a problem correctly after being given the explicit procedure. Additionally, one child in the right-blank condition who was coded as needing the explicit procedure did put the correct number in the blank for “$6 + 2 + 8 = 5 + _ _$” after receiving feedback only (the first level), but accompanied the solution with an incorrect explanation (“I added 6 and 5”) and proceeded to solve the next three problems incorrectly until being told the correct procedure.
correctly after being guided through the explicit procedure and then failed to implement a correct strategy on the final problem presented. So, generally once children understood, they maintained this learning in both conditions — at least during the intervention, when receiving consistent scaffolding as well as feedback about the correctness of their responses on each problem.

What is it about those kids who are able to learn with just a few levels of scaffolding compared to the kids who need more levels, or never are able to solve one problem correctly? I hypothesized that pretest encoding accuracy may be a factor here. I used overall pretest encoding accuracy in the following analyses; note, however, that conclusions are unchanged when only pretest encoding accuracy on the intervention problem format was considered. Using logistic regression, I examined whether pretest encoding accuracy predicted whether children were able to solve one problem correctly before being given any conceptual information (i.e., before the equal sign definition level), separately for each condition. Pretest encoding accuracy was not a significant predictor of solving at least one problem correctly prior to being given conceptual help in either the left-blank condition, $\hat{\beta} = 2.61$, $Wald \ (1, \ N = 36) = 2.99, \ p = .08$, or the right-blank condition, $\hat{\beta} = -0.69$, $Wald \ (1, \ N = 34) = 0.31, \ p = .58$. Finally, I examined whether pretest encoding accuracy mattered for whether children ever were able to learn from the instruction in each condition. Pretest encoding accuracy was not a significant predictor of solving at least one problem correctly during the intervention for either the left-blank condition, $\hat{\beta} = 3.17$, $Wald \ (1, \ N = 36) = 2.30, \ p = .13$, or the right-blank condition, $\hat{\beta} = 2.16$, $Wald \ (1, \ N = 34) = 1.09, \ p = .30$. 

63
Although pretest encoding accuracy was not a significant factor predicting children’s level of scaffolding needed to learn, it is possible that those children who needed only minimal scaffolding had more basic knowledge of equivalence outside of the symbolic context, and could have been more ready to learn from instruction. I did not measure children’s knowledge of equivalence in a concrete context at pretest, but I did measure children’s general math proficiency. Recall, however, that there were no differences among conditions in terms of general math proficiency at pretest, $p = .52$. It is possible however, that general math proficiency, regardless of condition, predicts children’s response to the intervention. In additional logistic regression analyses, I first examined whether pretest math proficiency (i.e., accuracy on the general math achievement items) predicted whether children were able to solve one problem correctly before being given any conceptual information. Pretest math proficiency was not a significant predictor of solving at least one problem correctly prior to being given conceptual help, $\hat{\beta} = 1.97$, $Wald (1, N = 70) = 3.32$, $p = .07$. Finally, I examined whether pretest math proficiency mattered for whether children ever were able to learn from the instruction. Pretest math proficiency was indeed a significant predictor of solving at least one problem correctly during the intervention, $\hat{\beta} = 3.26$, $Wald (1, N = 70) = 4.52$, $p = .03$.

3.2.5 Do Children Maintain and Transfer What They Learned in the Intervention?

Children who did not solve any problems correctly during the intervention continued to solve the rest of the problems all incorrectly (posttest, opposite format, and other formats), in both conditions. What about children who solved at least one intervention problem correctly ($n = 59$)? For these children, I calculated what percentage lost what they learned, maintained what they learned, or went beyond what they learned
to solve problems correctly of the opposite format. Criteria for being labeled as maintaining learning was solving the majority (i.e., at least two of three) of posttest problems correctly, whereas criteria for demonstrating knowledge beyond what was taught in the intervention was solving the majority (i.e., at least two of three) of problems of the opposite format correctly. Note that one child in each condition solved all three of the opposite format problems correctly but only one of the three posttest problems correctly; these children were each given credit for going beyond what they learned.4

For learners from the left-blank condition (n = 29), 14% lost what they learned, 24% maintained what they learned, and 62% went beyond what they learned. For learners from the right-blank condition (n = 30), 20% lost what they learned, 17% maintained what they learned, and 63% went beyond what they learned. I conducted a chi-square analysis to see if level of knowledge demonstrated after intervention differed by condition, and it was not significant, $\chi^2 = 0.74, p = .69$. Additionally, when I collapsed the maintained and beyond categories, a chi-square test examining only the categories of lost versus at least maintained was not significant, $\chi^2 = 0.40, p = .53$.

Furthermore, I examined correlations between the level of scaffolding needed to solve at least one problem correctly during the intervention and subsequent accuracy on the posttest and problems of the opposite format. The correlation between level of scaffolding needed and posttest accuracy was $\rho = -.24, p = .07$, and between level of

---

4 Additionally, one child in the right-blank condition solved all six problems after the intervention incorrectly (quickly with an add-to-equal-sign strategy), and realized when asked to explain his first solution that he solved them all incorrectly, and proceeded to change each to the correct number in the blank and explain them all with a correct, equalize strategy. For the purposes of accuracy reported above, I was strict and only gave credit for children’s responses prior to being asked to explain; however, in this case the child clearly was able to explain a correct strategy on problems that were both the intervention format and the opposite format, so he was classified in the beyond category (note that conclusions are unchanged, however, if he is classified as losing what he learned).
scaffolding and accuracy on problems of the opposite format was $\rho = -0.35, p = 0.01$. When correlations were examined separately by condition, the correlation between level of scaffolding needed to learn and accuracy on posttest problems was significant in the right-blank condition ($\rho = -0.36, p = 0.048$), but not in the left-blank condition ($\rho = -0.15, p = 0.44$), and the correlation between level of scaffolding needed to learn and accuracy on problems of the opposite format was also significant in the right-blank condition ($\rho = -0.46, p = 0.01$), but not in the left-blank condition ($\rho = -0.18, p = 0.35$). In other words, the fewer levels of scaffolding a child needed to solve at least one problem correctly in the intervention, the more accurate the child was after the intervention. This was especially strong for predicting accuracy on problems of the opposite type for children in the right-blank condition.

In a simpler analysis, the correlation between levels of scaffolding needed to solve at least one problem correctly and the level of future performance achieved (lost, maintained, or went beyond) for children in the right-blank condition was significant, $\rho = -0.45, p = 0.014$. However, the same correlation was not significant for children in the left-blank condition, $\rho = -0.21, p = 0.27$.

3.2.6 Task Difficulty Rating by Problem Format

Children’s ratings of task difficulty (on a 1-5 scale from very easy to very difficult) were normally distributed ($M = 3.14, SD = 1.10$). Overall, ratings of difficulty solving left-blank problems ($n = 53, M = 3.19, SD = 1.14$) were similar to ratings of difficulty solving right-blank problems ($n = 52, M = 3.10, SD = 1.07$), $p = 0.67$. Recall that for the experimental conditions, the rating of task difficulty was only administered after the last problem of the intervention, whereas for the control condition, it was after the
only problem given. I separated children who solved that final problem prior to the rating correctly \((n = 56)\) and children who solved it incorrectly \((n = 49)\) to examine ratings of task difficulty separately (note, however, that conclusions are unchanged when children are separated into those who ever solved one problem correctly, \(n = 59\), and those who did not, \(n = 46\)). For children who solved the final intervention problem correctly, ratings of difficulty solving left-blank problems \((n = 28, M = 3.36, SD = 0.87)\) were similar to ratings of difficulty solving right-blank problems \((n = 28, M = 3.11, SD = 1.03)\), \(p = .33\). For children who did not solve that problem correctly, ratings of difficulty solving left-blank problems \((n = 25, M = 3.00, SD = 1.38)\) were similar to ratings of difficulty solving right-blank problems \((n = 24, M = 3.08, SD = 1.14)\), \(p = .82\). Finally, when only children in the control condition (i.e., those who solved only one problem, without help, \(n = 35\)) were considered, the average rating of difficulty solving left-blank problems \((n = 17, M = 2.29, SD = 0.99)\) was not significantly different from the average rating of difficulty solving right-blank problems \((n = 18, M = 2.72, SD = 1.02)\), \(p = .22\).

3.2.7 Post-hoc Analyses

To test the hypothesis that children who solved right-blank problems experienced a greater cognitive load during the intervention than children who solved left-blank problems, I calculated the time it took children to solve each problem during the intervention for all conditions, as a proxy measure of cognitive load (cf. Sweller et al., 1998). Time was coded from when the experimenter finished saying the directions to when the child wrote a solution in the blank (note that if children decided to change their response, the time was only coded to their initial response). For the instances in which
children never provided a response (e.g., “I don’t know”) even after encouragement to try, the time was coded until the experimenter moved on to the next problem.

For the very first problem attempted in the experimental conditions, which was solved incorrectly by all participants, children in the right-blank condition spent, on average, a similar amount of time solving the problem ($M = 24.58$ s, $SD = 16.85$ s) as children in the left-blank condition did ($M = 24.29$ s, $SD = 15.22$ s), $p = .94$. When I calculated the average amount of time spent on every problem given for each participant, there was no difference in the amount of time spent solving each problem during the intervention between children in the right-blank condition ($M = 36.37$ s, $SD = 15.86$ s) and children in the left-blank condition ($M = 39.68$ s, $SD = 25.08$ s), $p = .52$. Note that when the problem on which the experimenter provided assistance in adding together the numbers on the left side of the equation (i.e., the explicit procedure level) is excluded, average solutions times remained similar between conditions, $p = .46$.

For children in the control condition, there was also no difference in time spent solving the assigned problem between those who were randomly assigned to solve a right-blank problem ($n = 18$, video malfunction for one child so analysis is out of 17 children) ($M = 37.24$ s, $SD = 29.27$ s) and those who were randomly assigned to solve a left-blank problem ($n = 17$; $M = 29.82$ s, $SD = 22.25$ s), $p = .41$. A closer examination by strategy use on this problem also demonstrated no statistically significant differences. Overall, solution times were highly variable, and I did not find evidence that problem format impacted the amount of time children spent solving the problems during the intervention.
3.3 Discussion

Results from Study 2 demonstrated that children in the right-blank condition outperformed children in the left-blank condition during the intervention, with a significantly higher percentage of children needing only minimal support to solve at least one problem correctly. Furthermore, the fewer levels of scaffolding children needed in the right-blank condition, the better they performed after the intervention, whereas that correlation was not present for children in the left-blank condition. Once children learned in each condition, they performed similarly on the posttest, opposite format and other format problems. Results suggest that it may be more efficient to teach math equivalence using right-blank versus left-blank problems.

3.3.1 Theoretical Implications

My experiment provided a critical test of competing accounts about the role of encoding in learning. From one account, it would be best to use problems that children are quite accurate at encoding (left-blank problems) during instruction because they are not as cognitively demanding (e.g., Sweller, 1994). Many practitioners and educational researchers believe the best environment for the learner is always one in which cognitive load is reduced (Sweller & Chandler, 1994). Learners have this bias as well, in which we generally perceive the ease of learning as reflecting the strength of knowledge we have constructed rather than appreciating the benefits of challenge during knowledge acquisition (Bjork & Bjork, 2011). Indeed, typically students and teachers measure the success of a given lesson based on how easy the information was to encode rather than on long-term retention (Bjork, 1994). Education researchers also advocate for simplifying...
encoding of material and eliminating extraneous information (e.g., Kaminski & Sloutsky, 2013; Kaminski, Sloutsky, & Heckler, 2008).

However, with desirable difficulties (Bjork, 1994), there is greater cognitive engagement with the material, which facilitates both translation of information into memory and subsequent retrieval (Craik & Tulving, 1975). My results were in line with this account, in that a greater percentage of children were able to generate a correct strategy with the right-blank intervention than the left-blank intervention. Previous research has demonstrated that children as much less accurate when encoding right-blank problems compared to left-blank problems (McNeil & Alibali, 2004), and my data was consistent with this, as children’s pretest encoding accuracy for right-blank problems was significantly lower than pretest encoding accuracy for left-blank problems. It may be that once we help children encode the problems correctly (e.g., providing perceptual support such as putting the equal sign in red), the extra features (number, plus sign) between the “=” and the “_” in a right-blank problem may induce cognitive conflict and lead children to create a new category in their mind for that type of problem because it cannot be assimilated into their existing arithmetic framework.

One interpretation of the change-resistance account would also predict it would be best to use left-blank problems during instruction, because the right-blank problems are more likely to activate children’s knowledge of the traditional arithmetic problem format (McNeil & Alibali, 2004, 2005b). However, my results (similar to results from Study 1) indicated that there might be more than one way to activate children’s knowledge of the traditional arithmetic problem format. Specifically, children’s run-on schema could be activated when presented with left-blank problems, suggesting that it would instead be
better to use right-blank problems during instruction. In fact, children are exhibiting change-resistance and adherence to the traditional arithmetic problem format in both cases. In the instance of right-blank problems, with the blank in the very last position, children’s adherence to the traditional arithmetic problem format results in the problem being misencoded as a traditional “a + b + c = ____” problem with the blank at the end. In the instance of left-blank problems, children’s adherence to the traditional arithmetic problem format does not result in incorrect encoding; instead, their adherence results in the problem being encoded correctly (the “= ____” chunk maps onto children’s expectation for where “=” and “___” appear in a math problem), but it is misinterpreted as a run-on traditional addition problem.

Overall, my results from Study 2 add to those from Study 1 and provide contrasting evidence to the widespread encoding-solving relation. Recall that the left-blank condition improved in their encoding of the problem format given during the intervention, whereas the right-blank condition did not. The prevailing view that there is a positive association between encoding and solving would suggest that these improvements in encoding should be accompanied by better solving performance at posttest than children in the condition which did not show encoding improvement. However, this was not the case. Children in both the left-blank and right-blank conditions performed similarly on the posttest, providing contrasting evidence to the view that encoding and solving are always positively associated.

The null effect for improvement in encoding of right-blank problems is similar in fact to a previous study that also failed to find improvements in encoding accuracy after a perceptual manipulation (equal sign in red) with right-blank problems only (Alibali,
Children are much less accurate when encoding right-blank problems than left-blank problems (McNeil & Alibali, 2004), and it may take much more than this brief intervention to get children to encode right-blank problems correctly. Note that the results from the present study differ from those reported by Alibali et al. (2009), in which children were able to improve their encoding accuracy after direct instruction on right-blank problems; however, the encoding representation assessment in that study included a recognition task as well as a reconstruction task at pretest and posttest, so students were both exposed to more right-blank problems and had more opportunities to demonstrate their knowledge. Recall that in the present study a greater percentage of children in the right-blank condition were able to generate a correct strategy prior to receiving conceptual instruction, but children’s pretest encoding accuracy was not a significant moderator of needing minimal support, or even of whether children ever solved a problem correctly during instruction. This adds further evidence that encoding and solving are not always linked, contrary to the prevailing account in the literature that encoding performance predicts solving performance (e.g., Kotovsky et al., 1985; Rittle-Johnson et al., 2001; Siegler, 1976).

Recall that about 25% of children in the right-blank condition were able to generate a correct strategy with minimal assistance (i.e., one or two levels of scaffolding), whereas none of the children in the left-blank condition were successful with such minimal support. This may be important, because these children in the right-blank condition were able to discover correct strategies to solve the problems without direct conceptual instruction, but simply with feedback on the correctness of their responses and having the equal sign presented in red. Findings from a recent meta-analysis suggest that
although unassisted discovery (i.e., no feedback, just exploration) does not benefit
learners, assisted discovery in which support such as feedback and scaffolding are
provided is more favorable for learning than explicit instruction (Alfieri, Brooks, Aldrich,
& Tenenbaum, 2011). In fact, some educators argue that children who acquire knowledge
on their own are more likely to transfer that knowledge than children who simply receive
direct instruction (e.g., Keith, Richter, & Naumann, 2010; McDaniel & Schlager, 1990).
In this way, generating the correct strategy with fewer levels of scaffolding (like more
children in the right-blank condition were able to do) could lead to better learning than
being given the conceptual instruction in order to solve the problem correctly (like more
children in the left-blank condition needed).

3.3.2 Practical Implications

Results provide educators with information about how children learn from
instruction on math equivalence. Specifically, the format in which teachers present
problems matters, and it may be easier for children to pick up the concept with right-
blank problems than with left-blank problems. However, it is important to note that I am
not advocating that children only be taught with a single format of problems. Rather,
research points to the benefits of interleaving problems of different formats for improving
learning (Kornell & Bjork, 2008; Rohrer, 2012; Rohrer, Dedrick, & Stershic, 2015;
Schmidt & Bjork, 1992). When a teacher or tutor has unlimited time to spend on teaching
children math equivalence, both right-blank and left-blank problems will be helpful to
incorporate, as well as problems with the blank prior to the equal sign (e.g., \( _ + a = b + c \)),
and problems presenting the equal sign outside the context of arithmetic (e.g., \( 28 = 28 \);
McNeil, 2008). Unfortunately, teachers often do not have unlimited time to spend on a
single concept during instruction and need to make decisions about which problems to work on with children in order to maximize efficiency. My results suggest that if teachers or tutors only have time to use one problem type in their instruction on math equivalence, they may benefit from using a right-blank problem to increase the likelihood that children will pick up the concept quickly.

Even with instruction set up to increase the likelihood that children will learn, misconceptions about math equivalence are difficult to overcome. Recall that in the present study, some children were never able to learn to solve at least one problem correctly in either condition, and many children either did not demonstrate that they maintained what they had learned on the posttest or were not able to transfer their knowledge to the opposite problem format. This is consistent with existing literature on children’s understanding of math equivalence, as some children fail to learn from interventions (e.g., Jacobs et al., 2007; Rittle-Johnson & Alibali, 1999), and others exhibit only shallow learning, failing to transfer to problems that differ in surface form (e.g., number of addends, presence of repeated addends) (e.g., Alibali, 1999; Alibali et al., 2009; Perry, 1991). Even when children do learn and transfer immediately following instruction on the problems, it is common for them to simply revert back to their ingrained misconceptions a few weeks after instruction (e.g., Cook et al., 2008; McNeil & Alibali, 2000).

3.3.3 Limitations and Future Directions

The present study is not without limitations. First, this experiment was only one session. It may be that differences in performance would emerge on a delayed posttest. Previous research has indicated many children revert back to their entrenched
misconceptions a few weeks after instruction (e.g., Cook et al., 2008; McNeil & Alibali, 2000). However, it is possible that because a greater percentage of children in the right-blank condition compared to the left-blank condition were able to discover the strategy on their own prior to any conceptual information being explicitly presented, children in the right-blank condition would be less likely to revert back to these misconceptions (Alfieri et al., 2011). Future research is needed to test this hypothesis.

Second, by the nature of the intervention setup, all children in the experimental conditions were given all five problems to solve. The fewer levels of scaffolding children needed to solve one problem correctly, the more opportunities they had during the rest of the intervention to solve problems alongside added conceptual instruction and receive feedback that they were solving the problems correctly. This more extensive positive feedback for a greater percentage of children in the right-blank condition could have increased the likelihood that children persisted in trying a certain strategy on the later assessments including problems of the opposite format and other formats (Deci, Koestner, & Ryan, 1999). Not only did children get more correct feedback, but also they had more opportunities for self-explanation of their correct strategies rather than incorrect strategies. Prompts for self-explanation have been shown to be effective tools for promoting learning and transfer (Chi, DeLeeuw, Chiu, & LaVancher, 1994; Rittle-Johnson, 2006; Siegler, 2002). This could be even more effective when children are solidifying their newly generated correct strategy along with feedback. Indeed, self-explanations have been found to be effective for resolving misconceptions (Allen, McNamara, & McCrudden, 2015). However, continued confirmation that one is solving problems correctly may not necessarily be beneficial, and could lead to overconfidence. It
may be that for children in the right-blank condition who had already generated a correct strategy, they were only passively processing the conceptual information the tutor provided on the rest of the intervention problems. Indeed, research has indicated the lack of struggle students experience can lead to superficial processing of information (Bjork, 1994; Wittwer & Renkl, 2008).

Third, in a similar manner, the nature of the intervention design and placement of the task difficulty rating at the very end of the intervention may have impacted results. I chose to include the rating only at the end of the intervention to keep the session feeling as natural as possible. However, the drawback is that children who ended up learning by the end may rate the problems as less difficult than they really thought they were. Indeed, a disadvantage of subjective ratings is that the participant’s frame of reference can change during the course of learning (Schnitz & Kürschner, 2007). The rating was included after the one problem solved in the control condition (randomly either a left-blank or a right-blank problem) in an effort to ascertain children’s feelings about the difficulty level independent of other influences, such as feedback from the experimenter. Future research could examine cognitive load using physiological measures or dual-task methodology in conjunction with subjective measures (see Schnitz & Kürschner, 2007 for a review).

Finally, with the data available from the present study I am unable to directly test the effect of children’s conceptual understanding on their performance during and after the intervention. Future research could assess conceptual understanding using measures of equal sign understanding in concrete contexts (e.g., Sherman & Bisanz, 2009), recognition of equations as true or false (e.g., Carpenter et al., 2003; Rittle-Johnson et al.,
2011), or explicit definitions of the equal sign (e.g., Baroody & Ginsburg, 1983; Behr et al., 1980) prior to a lesson with various symbolic problem formats to test specifically the impact of conceptual understanding on future problem solving of each format type. It may be that children with an emerging understanding of math equivalence as demonstrated on measures of conceptual understanding will require fewer levels of scaffolding to learn how to solve symbolic math equivalence problems correctly.
3.4 Figures

<table>
<thead>
<tr>
<th>How easy or difficult was it to solve all of those problems?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How easy or difficult was it to solve that problem?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy</td>
</tr>
</tbody>
</table>

Figure 3.1: Task difficulty rating for the experimental conditions (top panel) and control condition (bottom panel).
Figure 3.2: Levels of scaffolding.
Figure 3.3: Adjusted problem-solving accuracy across test phases by condition.
Figure 3.4: Unadjusted problem-solving accuracy by experimental condition (left-blank, $n = 36$; right-blank, $n = 34$) for intervention problems (top row) and posttest problems (bottom row).
Figure 3.5: Unadjusted problem-solving accuracy by experimental condition (left-blank, \( n = 36 \); right-blank, \( n = 34 \)) for opposite format problems (top row) and other format problems (bottom row).
Figure 3.6: Adjusted posttest encoding accuracy for each problem format by condition.
Figure 3.7: Unadjusted posttest encoding accuracy by condition (left-blank, $n = 36$; right-blank, $n = 34$; control, $n = 35$) for intervention (left column) and opposite problems (right column).
Figure 3.8: Percent of children learning at each level of scaffolding by condition.
TABLE 3.1
RANDOMIZED EXPERIMENT: PRETEST ASSESSMENT

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Set A</th>
<th>Set B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>301</td>
<td>301</td>
</tr>
<tr>
<td></td>
<td>−75</td>
<td>−75</td>
</tr>
<tr>
<td>2</td>
<td>The 16 stickers listed above are placed in a box. If one sticker is drawn from the box, which color is it most likely to be?</td>
<td>The 16 stickers listed above are placed in a box. If one sticker is drawn from the box, which color is it most likely to be?</td>
</tr>
<tr>
<td>3</td>
<td>$8 + 2 = _ + 6$</td>
<td>$8 + 2 = 6 + _$</td>
</tr>
</tbody>
</table>
TABLE 3.1 (CONTINUED)

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Set A</th>
<th>Set B</th>
</tr>
</thead>
<tbody>
<tr>
<td>What fraction of the group of umbrellas is closed?</td>
<td>What fraction of the group of umbrellas is closed?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><img src="image" alt="Umbrellas" /></td>
<td><img src="image" alt="Umbrellas" /></td>
</tr>
<tr>
<td>5</td>
<td>[ 7 + 5 + 4 = 7 + _ _ ]</td>
<td>[ 5 + 4 + 7 = _ _ + 7 ]</td>
</tr>
<tr>
<td>Encoding section</td>
<td>[ 7 + 1 = _ _ + 6 ]</td>
<td>[ 4 + 5 = 3 + _ _ ]</td>
</tr>
<tr>
<td></td>
<td>[ 4 + 5 = 3 + _ _ ]</td>
<td>[ 7 + 1 = _ _ + 6 ]</td>
</tr>
<tr>
<td></td>
<td>[ 8 + 3 + 9 + 8 = _ _ ]</td>
<td>[ 8 + 3 + 9 + 8 = _ _ ]</td>
</tr>
<tr>
<td></td>
<td>[ 3 + 5 + 4 = _ _ + 4 ]</td>
<td>[ 2 + 3 + 6 = 2 + _ _ ]</td>
</tr>
<tr>
<td></td>
<td>[ 2 + 3 + 6 = 2 + _ _ ]</td>
<td>[ 3 + 5 + 4 = _ _ + 4 ]</td>
</tr>
<tr>
<td>6</td>
<td>[ 5 + 4 + 7 = _ _ + 7 ]</td>
<td>[ 7 + 5 + 4 = 7 + _ _ ]</td>
</tr>
<tr>
<td>7</td>
<td>[ 5,631 + 286 = _ _ ]</td>
<td>[ 5,631 + 286 = _ _ _ _ ]</td>
</tr>
<tr>
<td>What shapes make up the faces of a square pyramid?</td>
<td>What shapes make up the faces of a square pyramid?</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td><img src="image" alt="Square Pyramid" /></td>
<td><img src="image" alt="Square Pyramid" /></td>
</tr>
<tr>
<td>9</td>
<td>[ 8 + 2 = 6 + _ _ ]</td>
<td>[ 8 + 2 = _ _ + 6 ]</td>
</tr>
<tr>
<td>10</td>
<td>[ 4 \times 3 = _ _ ]</td>
<td>[ 4 \times 3 = _ _ ]</td>
</tr>
</tbody>
</table>
### TABLE 3.2

**RANDOMIZED EXPERIMENT: POSTTEST, OPPOSITE FORMAT, AND OTHER FORMATS**

<table>
<thead>
<tr>
<th>Encoding</th>
<th>Posttest</th>
<th>Left-blank condition</th>
<th>Right-blank condition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Posttest 1</strong></td>
<td></td>
<td>$\begin{array}{c} 1 + 5 = _ _ _ + 2 \ 2 + 7 = 6 + _ _ _ \ 9 + 7 + 6 + 9 = _ _ _ \ 7 + 2 + 4 = _ _ _ + 4 \ 3 + 5 + 6 = 3 + _ _ _ \end{array}$</td>
<td>$\begin{array}{c} 2 + 7 = 6 + _ _ _ \ 2 + 7 = 6 + _ _ _ \ 9 + 7 + 6 + 9 = _ _ _ \ 3 + 5 + 6 = 3 + _ _ _ \ 7 + 2 + 4 = _ _ _ + 4 \ 3 + 5 + 6 = 3 + _ _ _ \ 1 + 5 = _ _ _ + 2 \ 2 + 7 = 6 + _ _ _ \ 9 + 7 + 6 + 9 = _ _ _ \ 7 + 2 + 4 = _ _ _ + 4 \ 3 + 5 + 6 = 3 + _ _ _ \end{array}$</td>
</tr>
<tr>
<td><strong>Posttest 2</strong></td>
<td></td>
<td>$\begin{array}{c} 6 + 8 + 2 = _ _ _ + 4 \ 7 + 4 + 9 = _ _ _ + 9 \ 9 + 6 + 5 = _ _ _ + 9 \ 2 + 9 + 7 = 5 + _ _ _ \ 6 + 5 + 9 = _ _ _ + 9 \end{array}$</td>
<td>$\begin{array}{c} 6 + 8 + 2 = _ _ _ + 4 \ 7 + 4 + 9 = _ _ _ + 9 \ 9 + 6 + 5 = _ _ _ + 9 \ 2 + 9 + 7 = 5 + _ _ _ \ 6 + 5 + 9 = _ _ _ + 9 \end{array}$</td>
</tr>
<tr>
<td><strong>Posttest 3</strong></td>
<td></td>
<td>$\begin{array}{c} 9 + 4 + 5 = _ _ _ + 7 \ 9 + 7 + 4 = _ _ _ + 9 \ 7 + 9 + 5 = _ _ _ + 8 \ 7 + 9 = 5 + _ _ _ + 8 \end{array}$</td>
<td>$\begin{array}{c} 9 + 4 + 5 = _ _ _ + 7 \ 9 + 7 + 4 = _ _ _ + 9 \ 7 + 9 = 5 + _ _ _ + 8 \end{array}$</td>
</tr>
<tr>
<td><strong>Opposite 1</strong></td>
<td></td>
<td>$\begin{array}{c} 7 + 5 + 6 = 3 + _ _ _ \ 9 + 6 + 5 = _ _ _ + 9 \ 7 + 9 + 5 = _ _ _ + 8 \ 7 + 9 = 5 + _ _ _ + 8 \end{array}$</td>
<td>$\begin{array}{c} 7 + 5 + 6 = 3 + _ _ _ \ 9 + 6 + 5 = _ _ _ + 9 \ 7 + 9 + 5 = _ _ _ + 8 \ 7 + 9 = 5 + _ _ _ + 8 \end{array}$</td>
</tr>
<tr>
<td><strong>Opposite 2</strong></td>
<td></td>
<td>$\begin{array}{c} 2 + 9 + 7 = 5 + _ _ _ \ 2 + 9 + 7 = 5 + _ _ _ \end{array}$</td>
<td>$\begin{array}{c} 2 + 9 + 7 = 5 + _ _ _ \end{array}$</td>
</tr>
<tr>
<td><strong>Opposite 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other 1</strong></td>
<td></td>
<td>$\begin{array}{c} 5 + _ _ _ + 6 = 9 + 8 + 2 \ _ _ _ + 5 = 7 + 8 + 3 \ _ _ _ + 5 = _ _ _ + 8 + 3 \end{array}$</td>
<td>$\begin{array}{c} 5 + _ _ _ + 6 = 9 + 8 + 2 \ _ _ _ + 5 = 7 + 8 + 3 \ _ _ _ + 5 = _ _ _ + 8 + 3 \end{array}$</td>
</tr>
<tr>
<td><strong>Other 2</strong></td>
<td></td>
<td>$\begin{array}{c} _ _ _ + 6 = 9 + 8 + 2 \ _ _ _ + 6 = 9 + 8 + 2 \end{array}$</td>
<td>$\begin{array}{c} _ _ _ + 6 = 9 + 8 + 2 \end{array}$</td>
</tr>
<tr>
<td><strong>Other 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 3.3
RANDOMIZED EXPERIMENT: INTERVENTION PROBLEMS

<table>
<thead>
<tr>
<th>Problem No.</th>
<th>Left-blank condition</th>
<th>Right-blank condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$6 + 2 + 8 = _ + 5$</td>
<td>$6 + 2 + 8 = 5 + _ $</td>
</tr>
<tr>
<td></td>
<td>(Given chance to re-solve if incorrect)</td>
<td>(Given chance to re-solve if incorrect)</td>
</tr>
<tr>
<td>2</td>
<td>$3 + 9 + 7 = _ + 5$</td>
<td>$3 + 9 + 7 = 5 + _ $</td>
</tr>
<tr>
<td>3</td>
<td>$8 + 7 + 5 = _ + 3$</td>
<td>$8 + 7 + 5 = 3 + _ $</td>
</tr>
<tr>
<td></td>
<td>(Given chance to re-solve if incorrect)</td>
<td>(Given chance to re-solve if incorrect)</td>
</tr>
<tr>
<td>4</td>
<td>$3 + 5 + 9 = _ + 7$</td>
<td>$3 + 5 + 9 = 7 + _ $</td>
</tr>
<tr>
<td></td>
<td>(Given chance to re-solve if incorrect)</td>
<td>(Given chance to re-solve if incorrect)</td>
</tr>
<tr>
<td>5</td>
<td>$2 + 6 + 7 = _ + 4$</td>
<td>$2 + 6 + 7 = 4 + _ $</td>
</tr>
</tbody>
</table>

*Children in the control condition were randomly assigned to receive one of these problems.
This project provides a critical test of one of the prevailing theories of children’s learning across domains, that poor encoding causes solving difficulties. Specifically, I found that the ease of encoding is not predictive of solving performance on math equivalence problems (demonstrated in Study 1) or ease of learning, as measured by the levels of scaffolding needed to solve at least one math equivalence problem correctly (demonstrated in Study 2). Furthermore, Study 1 demonstrated that encoding accuracy was correlated with solving accuracy on right-blank problems only, and Study 2 demonstrated that improvement in encoding accuracy on left-blank problems was not associated with better solving performance. Taken together, these results challenge the widespread assumption that encoding performance always predicts solving performance, and instead require that the prevailing view be expanded to include solvers’ interpretations of what they encode.

In an effort to impact children’s understanding of math equivalence before misconceptions become entrenched, my project provides educators with information about both the sources of children’s difficulties with math equivalence and how children learn from instruction on math equivalence. The change-resistance account of children’s difficulties with math equivalence (McNeil & Alibali, 2005b) suggests that children’s performance and learning will be hindered when their knowledge of traditional arithmetic
is activated. One interpretation of the change-resistance account suggests that the right-blank problem format should be worse in terms of performance and learning because this format has been shown to activate children’s knowledge of the traditional problem format and be encoded inaccurately (McNeil & Alibali, 2004). However, the current project indicates that left-blank problems can activate children’s knowledge of the traditional problem format when they are encoded accurately. Instead of misencoding the problems, children interpret the correctly encoded problem as a run-on traditional addition problem so it is consistent with their established knowledge of traditional arithmetic. In this way, children’s knowledge of traditional arithmetic hinders processing of both math equivalence problem formats. In the case of right-blank problems, in which the position of the blank is at the end of the problem, children’s knowledge of traditional arithmetic leads them to misencode the problem as a traditional “c + b + c = ___” problem with the blank in the final position. In the case of left-blank problems, children’s knowledge of traditional arithmetic leads them to misinterpret correctly encoded problems as run-on traditional addition problems that require multiple steps of traditional computation.

Children’s traditional arithmetic instruction in the U.S. is thought to contribute to difficulties solving math equivalence problems, in which adding up the numbers left to right does not result in success (McNeil & Alibali, 2005b). Results from my project provided additional support for this in that children’s run-on traditional arithmetic schema seemed to underlie difficulties with left-blank problems, in which children will add the numbers to the left of the equal sign, put the sum in the blank, and then view the rest of the problem as a subsequent calculation. This overwhelming left-to-right tendency in problem solving is in line with work by McCrink and colleagues on operational
momentum (McCrink, Dehaene, & Dehaene-Lambertz, 2007) and work by others on spatial-numerical associations in which large amounts are expected to be on the right side rather than the left side of a display (e.g., Dehaene, Bossini, & Giraux, 1993; Fischer, Castel, Dodd, & Pratt, 2003). Children’s mathematics input in the U.S. is inundated with traditional equation formats like $a + b + c = \_\_\_$ (e.g., McNeil et al., 2006; Powell, 2012; Seo & Ginsburg, 2003) and number lines increasing from left to right, and children even read English from left to right. Building from the change-resistance account (McNeil & Alibali, 2005b), a greater effort by educators to increase the variability of input children receive in terms of problem formats is likely to lessen the negative impact of this left-to-right schema (e.g., McNeil, Fuhs, Keultjes, & Gibson, 2011; McNeil et al., 2011). Indeed, for learners to extract invariant structure (in this case, to understand that “=” always means that two numbers or mathematical expressions are equivalent), that item (“=”) must appear in a variety of contexts. Once children are fluent with a variety of problem formats, they will have a greater capacity to focus on new concepts to be learned after their cognitive load is reduced (Kellman et al., 2008).

Results not only add to the literature on children’s understanding of math equivalence, a critical concept for later mathematics understanding, but they also add to the general body of research about the role of encoding in performance and learning. Although in this project I found that the interpretation of what is encoded matters more than perceptual encoding, this does not discount the importance of perceptual factors in mathematical reasoning. We know that spatial factors can affect implementation of order of operations rules when calculating (Kirshner, 1989; Landy & Goldstone, 2007, 2010), and even conceptual interpretations of equations (Jiang, Cooper, & Alibali, 2014). A
large body of research highlights the benefit of supporting learners’ perceptual encoding in domains such as fractions, algebra, measurement, and arithmetic (e.g., Goldstone, Landy, & Son, 2010; Kellman et al., 2008; Kellman, Massey, & Son, 2010; Prather & Alibali, 2011). Rather, my results suggest children need support to understand what the symbols which they are encoding mean before they can interpret the mathematical structure. This is in line with Booth’s (1989) statement that “algebraic representation and symbol manipulation…should proceed from an understanding of the semantics or referential meanings that underlie it” (p. 58). In all domains, it is important for learners to understand the underlying concepts so that they can recognize the range of problems for which a particular procedure applies regardless of specific surface features (Siegler, 1996).

Proficiency in mathematics, especially algebra, depends on the ability to accurately and flexibly perceive the structure of mathematical expressions (Kellman et al., 2010; Linchevski & Livneh, 1999; Lüken, 2012). Indeed, Kieran (2007) notes that a structural understanding of algebra stems from students’ ability to “see” the abstract ideas that are hidden behind symbols. The results of my research indicate that the positive association between encoding and solving is not universal, and instead this view needs to be expanded to include solvers’ interpretations of particular problem features. Math is a language of symbols, but students who learn to simply apply procedures to symbols are missing the point. Deep conceptual knowledge about what those symbols mean, in particular the relational nature of the equal sign, along with accurate perception of equations, may propel students to fare better in their transition to algebra (Alibali et al., 2007; Knuth et al., 2006).
More broadly, my results add to the large body of literature that suggests that small changes to the environment can have an impact on both performance and learning. This is seen in a variety of contexts, from problem orientation (McNeil et al., 2011; Trbovich & LeFevre, 2003), to perceptual richness (Kaminski & Sloutsky, 2013; McNeil, Uttal, Jarvin, & Sternberg, 2009; Petersen & McNeil, 2013), to the level of concreteness of instruction (Fyfe, McNeil, & Borjas, 2015). Adding to this work, my project demonstrates that a small perceptual change of math equivalence problem format (right-blank versus left-blank) matters for both solving performance and how quickly children are able to learn.

4.1 Conclusion

The present findings are in contrast to the widely-held theory that children’s encoding performance always predicts solving performance, and suggest that in certain domains it may be that the interpretation of what is encoded is more important than accurate encoding. These findings are of theoretical and practical importance, as perceptual encoding is necessary for interpreting every mathematical expression encountered. Beyond perceptual encoding, results suggest a need to focus more on understanding what mathematical symbols mean before effective problem-solving procedures can be applied.
APPENDIX A:
RANDOMIZED EXPERIMENT: SESSION SCRIPT

PRETEST: SOLVING

Make sure the camera is recording with a good view of the child and their workspace. You should be sitting down with the child, diagonally across from them.

Say: We’re going to do some different activities together today. Sometimes you’ll solve some math problems at the table, sometimes we’ll do a memory-type activity here, and sometimes we’ll solve problems at the board. OK?

Put the pretest packet (#1A or #1B) and pen in front of the child.

I’m going to show you some math problems. This isn’t a test, so don’t worry if you aren’t sure how to solve them. I’m most interested to see how you think about the problems, so I want you to try your best on each one.

Say: Here’s the first problem. (turn page). Write your response on the paper.

*If the child seems to be confused, or if the child is taking a very long time, say:

*Say: Remember, I’m interested to see how you think about the problems, so just solve it as best as you can. I want to know how YOU think about the problems.

After child solves problem:

Say: Let’s try the next one (turn page for child). Read instructions.

Repeat for each problem, pointing to blank whenever blank is mentioned.

*If child says response aloud, say: Please write your response on the paper.

*If child gets stuck on large addition problem, can say: You can do any work on the page (point) that will help you solve the problem.

At first encoding page (E1), transition to encoding directions on following page after saying: OK, now let me tell you the directions for this part. (Stand up).
PRETEST: ENCODING

You will need 5 laminated equations (labeled PRE). YOU SHOULD STAND FOR THIS ACTIVITY. These equations should always be presented in the ASSIGNED order (A/B):

All problems should be placed face down in a pile on your chair. Take problems one at a time from the pile. After each problem is shown, place it face down on the floor.

Say: **In this task, I’m going to show you some math problems. This time you don’t have to solve the problems. Instead, you just have to remember what you see and write it on here** (gesture to child’s blank page where he/she will write the first problem). *So, this is what we’ll do: I’ll hold up a problem for five seconds (hold hands like you are holding up a problem), and **AFTER I put the problem down** (put hands down), you’ll write exactly what you saw. OK?* (Wait for nod). **Remember, you don’t have to solve the problems; you just need to write exactly what you see AFTER I put the problem down.** (Pick up first problem, and hold it face down toward your body).

*Say: **Here’s the first** (second, third, etc.) **problem. Ready?** (Wait for nod).

*Flip the card up for the child to see for 5 seconds (count to yourself: 1 cognition, 2 cognition, 3 cognition, 4 cognition, 5 cognition); then put the problem on the floor face down to hide the card from the child.

*If child starts writing early, say: **Remember to wait until after I put the problem down before you start writing.** (Have them write on back of that page if they made marks on page).

*Say: **OK, Write exactly what you saw.**
*Say: **Flip to the next page.** (Make sure the child flips to the next blank page.)

Repeat the * steps above for the next 4 problems.

Say: **OK, we’re finished with that part. Now you’ll solve a few more problems in this packet.** CONTINUE READING INSTRUCTIONS ON EACH PAGE TO END.

After child finishes final problem:

Say: **OK, we’re finished with that task. Great job!**

<<< LET CHILD CHOOSE STICKER >>>>
Intervention: LB or RB

EQUATION SOLVING

You will need 5 laminated equations (labeled by condition and numbered 1-5). Use the clip (or magnet) to mount the equation on the dry erase board. All problems should be placed face down in a pile on the floor. Take problems one at a time from the top of the pile. After each problem is solved, place the solved problem face down on the floor. Before starting, make sure the camera is recording, and make sure it has a good view of the problem. After clipping each problem to the easel, STAND NEXT TO THE RIGHT SIDE OF THE EASEL (when facing it) for all parts of the equation solving script.

**We want to see the problem on the video even when the child is standing in front of it, but don’t place the problem too high. It should be just above the child’s head.

Say: I’m going to show you some math problems one at a time. You might not have seen problems like these before, and that’s OK. I’m most interested to see how you think about the problems. It’s not a test, so don’t worry if you aren’t solving them the right way. I just want you to try your best.

PUT UP PROBLEM #1. $6 + 2 + 8 = \_ \_ + 5$ or $6 + 2 + 8 = 5 + \_ \_ $

Say: Try your best to solve the problem, and then write the number that goes in the blank (point to the blank [with your index finger]; hand child marker).

*If the child seems to be confused, or if the child is taking a very long time, say:

*Say: Remember, I’m interested to see how you think about the problems, so just solve it as best as you can, and put the number in the blank. I want to know how YOU think about the problems.

*If child still unable to provide a response after multiple encouragement prompts, *Say: Can you tell me how you were thinking about solving that problem?

Wait until the child has finished solving the problem. Take the marker as you--

Say: Can you tell me how you got $x$? ($x =$ child’s solution) (Point to the solution [with your index finger] when you say $x$.) Wait for response.

*If the child says “I counted”, “I added” or “I used my fingers”, say:

*Say: Can you tell me which ones you counted (added)? (While you say this, gesture with your hand from left to right just below the problem).

**NOTE: PROMPTS WITH * ABOVE ARE TO BE USED THROUGHOUT AS NEEDED.
* If the child expresses a correct strategy, say:
Great! That is a correct way to solve this problem. Let’s keep thinking about these problems.

>> GO TO PROBLEM #2

*If the child expresses an incorrect strategy FIRST TIME, say:
Well, that’s a really good try, but it’s not a correct way to solve the problem (erase solution), but that’s okay. These problems are challenging for kids your age. And we can try again, okay? Can you think of another way to solve the problem?

>> TRY PROBLEM 1 AGAIN

Say: Try again to solve the problem, and then write the number that goes in the blank (point to the blank [with your index finger]; hand child marker).

Wait until the child has finished solving the problem. TAKE THE MARKER as you--

Say: Can you tell me how you got x? (Point to the solution). Wait for response.

Provide feedback with correct strategy prompt from above or:

*If the child expresses an incorrect strategy AFTER FIRST TIME, say:
Well, that’s a really good try, but it’s not a correct way to solve the problem (erase solution), but we can try again, okay? Let’s keep thinking about these problems.

*If child expresses a different incorrect strategy than on first (or any) attempt, say:
Well, that’s a really good try, but it’s not a correct way to solve the problem (erase solution), but we can try again, okay? I really like that you’re trying to think of new ways to solve these problems. Let’s keep thinking about these problems.

*If child is discouraged at any point, say: Remember, this isn’t a test. You’re doing a great job of telling your ideas about these problems.

PUT UP PROBLEM #2.  $3 + 9 + 7 = \_ \_ + 5$  or  $3 + 9 + 7 = 5 + \_ \_ $

Say: Let’s see if this helps you think about the problem. Try your best to solve the problem, and then write the number that goes in the blank (point to the blank [with your index finger]; hand child marker).

*If child asks something like “Why is the equal sign red?,” say:
I’m interested in how you think about the problem when the = is red.

Wait until the child has finished solving the problem. TAKE THE MARKER as you--

Say: Can you tell me how you got x? (Point to the solution). Wait for response.

Provide feedback with prompts from above.
PUT UP PROBLEM #3. \( 8 + 7 + 5 = \_ \_ + 3 \) or \( 8 + 7 + 5 = 3 + \_ \_ \)

Say: Let’s see if this helps you think about the problem. Make sure you notice where the equal sign is in this problem. Can you point to the equal sign? (child points)

*If child is unsure or incorrect: This is where the equal sign is (point).
   Can you point to the equal sign? (child points).
*If child is correct: Great!

*Repeat prompts to have child solve problem (Now try your best…) and provide feedback as above.

*If correct >> After feedback, GO TO PROBLEM #4

*If incorrect >> After feedback, then:
   (Put marker down so your hands are free to gesture).

Say: Can you point again to the equal sign in this problem? (child points). Great! What does the equal sign mean? (allow time for child to respond, providing encouragement as needed).

*If child provides relational definition, say: That’s right, the equal sign means that what is on one side of it (gesture circle around L side with L hand) has to be the same amount as what is on the other side of it (gesture circle around R side with R hand).

*If child provides any sort of non-relational definition, say: You know, a lot of kids think that’s what the equal sign means. But actually, it doesn’t mean that! The equal sign means that what is on one side of it (gesture circle around L side with L hand) has to be the same amount as what is on the other side of it (gesture circle around R side with R hand).

*If child does not provide response, say: The equal sign means that what is on one side of it (gesture circle around L side with L hand) has to be the same amount as what is on the other side of it (gesture circle around R side with R hand).

*Repeat prompts to have child solve problem (Try again to solve…) and provide feedback as above.
PUT UP PROBLEM #4.  $3 + 5 + 9 = \_ \_ \_ + 7$ or $3 + 5 + 9 = 7 + \_ \_ \_ \$
(Put marker down so your hands are free to gesture).

Say: Let’s see if this helps you think about the problem. Can you point to the equal sign in this problem? (child points). What does the equal sign mean? (child responds).

*Repeat equal sign definition response prompts as above.

**NOTE: If child gave COR solution to #3 they haven’t yet answered definition question.

*If child did not provide non-relational definition, add “Remember, the equal means that…” prompt and say:
   So, what does the equal sign mean? (child responds). The equal sign means that what is on one side of it (gesture circle around L side with L hand) has to be the same amount as what is on the other side of it (gesture circle around R side with R hand).

*NOTE: If child has done EQ/AS, language below will be “you knew that…”

Say: So, because there is an equal sign (point to =), the amount on this side of the equal sign (gesture circle around L side with L hand) has to be the same as the amount on this side of the equal sign (gesture circle around R side with R hand). So, you need to find a number to go in the blank (point to blank) so that this side (gesture circle around L side with L hand) has the same amount as this side (gesture circle around R side with R hand).

*Repeat prompts to have child solve problem (Try your best…) and provide feedback.

*If correct >> After feedback, GO TO PROBLEM #5

*If incorrect >> After feedback, then:
   (Put marker down so your hands are free to gesture).

Say: Can you point to the equal sign again in this problem? (child points).

Say: So, because there is an equal sign (point to =), the amount on this side of the equal sign (gesture circle around L side with L hand) has to be the same as the amount on this side of the equal sign (gesture circle around R side with R hand). So, we need to find a number to go in the blank (point to blank) so that this side (gesture circle around L side with L hand) has the same amount as this side (gesture circle around R side with R hand).

Say: Let’s solve this problem. First, what is $3 + 5 + 9$? (point to each number). Allow time for child to add; provide feedback until child reaches correct sum. (e.g., So which numbers did you add together first?)
Say: **Right, so this side** (circle gesture around L side with L hand) **has 17.**

FOR LEFT BLANK CONDITION:

**Okay, so what number** (point to blank with R index finger) **plus** (point) **7** (point) **is equal to 17** (gesture around L side with L hand)? **That number should go in the blank**. (point to blank with R index finger; hand child marker).

FOR RIGHT BLANK CONDITION:

**Okay, so 7** (point to 7 with R index finger) **plus** (point) **what number** (point to blank) **is equal to 17** (gesture around L side with L hand)? **That number should go in the blank**. (point to blank with R index finger; hand child marker).

*Can repeat procedure once if needed (after other encouragement prompts)*

*After child writes solution, provide feedback:

**Great, that is a correct way to solve the problem. Let’s try another one.**

*If the child expresses an incorrect strategy, say:

**Well, that’s a really good try, but it’s not a correct way to solve the problem** (erase solution). **Let’s try another one.**

**PUT UP PROBLEM #5. 2 + 6 + 7 = ___ + 4** or **2 + 6 + 7 = 4 + ___**

Say: **Try your best to solve the problem, and then write the number that goes in the blank** (point to the blank [with your index finger]; hand child marker).

Wait until the child has finished solving the problem. **TAKE THE MARKER** as you--

Say: **Can you tell me how you got x?** (x = child’s solution) (Point to the solution [with your index finger] when you say x.) Wait for response.

Say: **OK, we’re done with those problems. Great job! You can have a seat now.**
TASK DIFFICULTY QUESTION: LB or RB

Say: I’m interested in what you think about the problems you just solved. I’m going to read a question to you and I want you to circle the word that best tells what you think.

Put *LB/RB* rating question in front of child.

Say: How easy or difficult was it to solve all of those problems?

Say: Was it very easy (circle), easy (circle), not easy or difficult (circle), difficult (circle), or very difficult (circle)? Circle the word that best tells what you think (sweep back and forth). Give child pen. Wait for child to circle.

*If child circles more than one response, say: Please circle the one that best tells what you think. Can you pick just one?

Say: OK, we’re finished with that part. You’re doing a great job.

<<< LET CHILD CHOOSE STICKER >>>>
**Intervention: Control**

**WORD SEARCH**

Give child word search.

Say:  **Try your best to find as many words as you can in this word search. After a few minutes, I’m going to tell you about some strategies to help you find the words.** Give child pen.

Start stopwatch.

*NOTE: If child asks at any point whether there are any backwards words, there are not! They are left to right, top to bottom, or diagonal, but none are backward.*

After 2 minutes, let stopwatch continue but start scaffolding:

*If child is clearly struggling before 2-minute mark, begin scaffolding.
*If child insists that they do not want help, let them continue working alone but that you can help if they get stuck.

Say:  **Great job! Ok, let’s talk about some strategies for completing word searches. First, it’s helpful to cross off the words in the list as you find them.** (If child did this, “Great job!”; If child did not do this, assist child in crossing off any words they have found).

Say:  **Let’s try to find some more words.**

*If child has not found “football” yet: “Football” might be an easy one to find because it has double letters in it. What double letters do you see in the word “football”? (child says “oo” or “ll,” ask for other double letter if they missed it). Let’s see if we can find those double letters in the word search. That could help us find the word “football.” Let child search, assist as needed – “football” is straight down in the bottom left corner.*

Say:  **One of the tricks that helps you find a certain word you are looking for is to look for the first letter of that word and then do a circle search around it to see if any of the letters next to it are the second letter in that word. Pick one of these words to look for, and try that strategy.** Assist child as appropriate.

*If child completes word search early, give child word scramble to work on. *If child struggling, encourage them to try, if needed bring out word scramble.

After stopwatch hits 10 minutes total:

Say:  **Okay, it’s time to move on to our next task. Great job with that!** Ask child to put paper under their chair, tell them they can finish it at home.
EQUATION SOLVING: CONTROL

NOTE: YOU WILL ADMINISTER ONLY ONE PROBLEM: #5 LB or #5 RB.

Use the clip (or magnet) to mount the equation on the dry erase board. After the problem is solved, place it face down on the floor. Before starting, make sure the camera is recording, and make sure it has a good view of the problem. After clipping each problem to the easel, STAND NEXT TO THE RIGHT SIDE OF THE EASEL (when facing it) for all parts of the equation solving script.

**We want to see the problem on the video even when the child is standing in front of it, but don’t place the problem too high. It should be just above the child’s head.

Say:  **Now I’m going to show you a math problem. You might not have seen a problem like this before, and that’s OK. I’m most interested to see how you think about the problem. I want you to try your best to solve the problem and then use this marker (show marker) to write a number in the blank. Do you understand the directions?** (Take cap off marker and hand marker to the child.)

Clip problem #5 (LB or RB, depending on random assignment) on the easel.

*Say:  **Try your best to solve the problem, and then write the number that goes in the blank** (point to the blank [with your index finger]).

*If the child seems to be confused, or if the child is taking a very long time, say:

*Say:  **Remember, I’m interested to see how you think about the problems, so just solve it as best as you can, and put the number in the blank. I want to know how YOU think about the problems.**

*Wait until the child has finished solving the problem. **TAKE THE MARKER** as you--

*Say:  Can you tell me how you got x? (x = child’s solution) (Point to the solution [with your index finger] when you say x.) Wait for response.

*If the child responds with “I counted”, “I added” or “I used my fingers”:

*Say:  Can you tell me which ones you counted (added)? (While you say this, gesture with your hand from left to right just below the problem.)

Say:  **OK, we’re done with that part. Great job! You can have a seat now.**
TASK DIFFICULTY QUESTION: CONTROL

Say: I’m interested in what you think about the problem you just solved. I’m going to read a question to you and I want you to circle the word that best tells what you think.

Put *CONTROL* rating question in front of child.

Say: How easy or difficult was it to solve that problem?

Say: Was it very easy (circle), easy (circle), not easy or difficult (circle), difficult (circle), or very difficult (circle)? Circle the word that best tells what you think (sweep back and forth). Give child pen. Wait for child to circle.

*If child circles more than one response, say: Please circle the one that best tells what you think. Can you pick just one?

Say: OK, we’re finished with that part. You’re doing a great job.

<<< LET CHILD CHOOSE STICKER >>>>
POSTTEST: ENCODING

You will need 5 laminated equations (labeled POST). YOU SHOULD STAND FOR THIS ACTIVITY. These equations should always be presented in the ASSIGNED order (A/B):

All problems should be placed face down in a pile on your chair. Take problems one at a time from the top of the pile. After each problem is shown, place it face down on the floor. You will need a spiral notecard booklet and a pen (for the child). Write the child’s ID number on the first blank page, and then flip to the next blank page. Before starting, make sure the camera is recording, and make sure it has a good view of the child.

Say: This is just like the activity we did earlier. In this task, I’m going to show you some math problems. This time you don’t have to solve the problems. Instead, you just have to remember what you see and write it on here (give the child the notecard booklet and pen and point to the spot where he/she will write). So, this is what we’ll do: I’ll hold up a problem for five seconds (hold hands like you are holding up a problem), and AFTER I put the problem down (put hands down), you’ll write exactly what you saw. OK? (Wait for nod). (Pick up first problem, and hold it face down toward your body).

*Say: Here’s the first (second, third, etc.) problem. Ready? (Wait for nod).

*Flip the card up for the child to see for 5 seconds (count to yourself: 1 cognition, 2 cognition, 3 cognition, 4 cognition, 5 cognition); then put the problem on the floor face down to hide the card from the child.

*If child starts writing early, say: Remember to wait until after I put the problem down before you start writing. (Have them flip to next notecard if they made marks on page).

*Say: OK. Write exactly what you saw.

*Say: Flip to the next page. (Make sure the child flips to the next blank notecard.)

Repeat the * steps above for the next 4 problems.

Say: OK, we’re finished with that part. You’re doing such a good job. Let’s move on to the next task.
POSTTEST AND OPPOSITE FORMAT TEST

Sit down with child (diagonally across from them). Give child packet #2L or #2R (depending on condition), opened to first page.

Say: Try your best to solve the problem, and then write the number that goes in the blank (point to blank).

*If the child seems to be confused, or if the child is taking a very long time, say:

*Say: Remember, I’m interested to see how you think about the problems, so just solve it as best as you can, and put the number in the blank. I want to know how YOU think about the problems.

After child solves problem, flip the page and read instructions for next one, until end.

After child finishes solving all problems, TAKE PEN and go back to the first one:

Ask: Can you tell me how you got X?  (Point to the solution when you say x.) (Child explains)

*REPEAT FOR EACH PROBLEM*

*If the child responds with “I counted”, “I added” or “I used my fingers”, say:

*Say: Can you tell me which ones you counted (added)? (While you say this, gesture with your hand from left to right just below the problem.)

_EXTRA P R O M P T  I F N E E D E D  F O R  6 th  P R O B L E M  O N L Y!_ 

If the child doesn’t show evidence of using grouping in their explanation (when COR):

Say: Is there a way to solve this problem without having to add all these numbers (gesture to left side)?

Say: OK, great job with those, now we’ve just got a few more tasks for today.
PROBLEMS OF OTHER FORMATS

Put packet (#3A or #3B) in front of child, opened to first page.

Say: Try your best to solve the problem, and then write the number that goes in the blank (point to blank). Give child pen.

*If the child seems to be confused, or if the child is taking a very long time, say:

*Say: Remember, I’m interested to see how you think about the problems, so just solve it as best as you can, and put the number in the blank. I want to know how YOU think about the problems.

After child solves problem, flip the page and read instructions for next one, until end.

After child finishes solving all problems, TAKE PEN and go back to the first one:

Ask: Can you tell me how you got $X$? (Point to the solution when you say $x$.) (Child explains)

*REPEAT FOR EACH PROBLEM*

*If the child responds with “I counted”, “I added” or “I used my fingers”, say:

*Say: Can you tell me which ones you counted (added)? (While you say this, gesture with your hand from left to right just below the problem.)

Say: OK, we’re all done with that. Great job!

LESSON

None or few correct: Basic lesson with blank type solved incorrectly
Most correct: Teach other types of problems solved incorrectly
All correct: Teach grouping strategy; if child picks up quickly, teach Relational

After Lesson:

Say: We’re finished with all of our activities today. I know I asked you to do a lot of things for me today. You did such a good job!! And I really appreciate you working so hard. Now, I have a few special gifts to give you to say thanks for all your hard work!

<<<< LET CHILD CHOOSE PENCIL & PENCIL TOP >>>>
Asterisks indicate studies included in the integrative data analysis.


