GESTURING MAY NOT ALWAYS MAKE LEARNING LAST

A Thesis

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

Master of Arts

by

Caroline E. Byrd

Nicole M. McNeil, Director

Graduate Program in Psychology
Notre Dame, Indiana
April 2014
© Copyright 2014

Caroline E. Byrd
GESTURING MAY NOT ALWAYS MAKE LEARNING LAST

Abstract

by

Caroline E. Byrd

Studies suggest that mimicking specific gestures prior to math instruction facilitates learning. However, benefits could be due to the eye movements that accompany gesture, rather than to gesture per se. Children (M age = 8 yrs, 9 mos) who solved pretest equations incorrectly were taught a correct strategy for solving equations. They were randomly assigned to mimic gestures instantiating the strategy, the eye movements that accompany those gestures, or speech only prior to and during instruction. Children completed an immediate posttest and a 4-week follow-up test. I hypothesized that children in the eye movement and gesture conditions would retain more from instruction when compared to children in the speech only condition. Posttest performance was similar across conditions. Contrary to hypotheses, children in the gesture condition retained less from instruction when compared to children in the other conditions. Results suggest that there may not always be benefits of gesture during instruction.
This is for my parents, Rob and Shirley.
ACKNOWLEDGMENTS

This research was funded by a National Science Foundation Graduate Research Fellowship and a Notre Dame Institute for Scholarship in the Liberal Arts research grant.

Thanks to my advisor, Nicole McNeil, as well as Sidney D’Mello, Susan Wagner Cook, Jill Lany, James Brockmole, Gretchen Busl, and members of the Cognition Learning and Development Lab at Notre Dame. Thanks also to the students, parents, teachers, and administrators at the participating school and afterschool programs.
CHAPTER 1:
INTRODUCTION

Students have to take and pass algebra if they want to attend college. Indeed, children who struggle with early algebra are not likely to gain the upper-level mathematics knowledge needed for many professional positions in the workplace (National Mathematics Advisory Panel, 2008). Unfortunately, first year algebra courses have been described as an “unmitigated disaster” for most students in the United States (National Research Council, 1998, p.1). Lack of readiness for algebra can be traced back to misunderstandings of pre-algebra concepts in elementary school (Carpenter, Franke, & Levi, 2003; Falkner, Levi, & Carpenter, 1999; Knuth, Stephens, McNeil, & Alibali, 2006). For example, most children (ages 7 to 11) in the U.S. do not understand how to solve *math equivalence problems*, which are equations that have operations on both sides of the equal sign (e.g., $3 + 4 + 5 = 3 + ___$; Alibali, 1999; McNeil, 2005, 2008; McNeil & Alibali, 2005b; Perry, 1991; Perry, Church, & Goldin-Meadow, 1988; Rittle-Johnson & Alibali, 1999). Difficulties with these problems are not easily “fixed” by instruction, as children often revert back to old, incorrect strategies a few weeks after being taught a correct strategy (Cook, Mitchell, & Goldin-Meadow, 2008).

When children attempt to solve math equivalence problems, they typically use a variety of strategies, both incorrect and correct. Some examples of children’s incorrect
strategies are adding up all the numbers in the problem, adding the numbers that appear to the left of the equal sign, or simply taking one number from the left side of the equation and putting it in the blank (Alibali & Goldin-Meadow, 1993; Perry et al., 1988). Although most children do not correctly solve math equivalence problems prior to instruction, many children are able to learn a correct procedure from a lesson on math equivalence (Perry et al., 1988; Rittle-Johnson & Alibali, 1999), at least temporarily. Various correct strategies children may use to solve math equivalence problems (e.g., \(a + b + c = a + __\)) include grouping the two numbers together that do not appear on both sides of the equation, adding the numbers on the left side of the equation and subtracting the number on the right, or figuring out what number would make both sides of the equation sum to the same total (Alibali & Goldin-Meadow, 1993; Perry et al., 1988).

While children are explaining their solutions to math equivalence problems, many children gesture spontaneously, typically in ways that convey particular procedures for solving the problem. Often, the procedure children convey in gesture is the same as (that is, “matches”) the procedure they convey in the speech accompanying that gesture. For example, for the problem \(3 + 4 + 5 = 3 + __\), a given child is noted as conveying a single procedure if the child indicates that he or she added all the numbers in the problem, both in speech (“I added 3 plus 4 plus 5 plus 3 and I got 15”) and in gesture (child points at the left 3, 4, 5, right 3, and the blank) The gestures produced by children during their explanations of solutions to math equivalence problems, however, do not always convey the same procedure as the speech that accompanies that gesture.
For example, if a child, in speech, indicates that he or she added only the numbers on the left side of the problem (“I added 3 plus 4 plus 5”) but, in gesture, indicates that he or she considered all of the numbers in the problem (child points at left 3, 4, 5, right 3, and the blank), this child is noted as conveying two procedures (i.e., a gesture-speech mismatch). Research has shown that children who produce many gesture-speech mismatches in their explanations of math equivalence problems benefit more from instruction than children who produce few gesture-speech mismatches (Perry et al., 1988). Researchers have also found direct evidence that gesture is a vehicle through which children express knowledge, and they are able to access knowledge they uniquely expressed in gesture in a different modality on other tasks (Garber, Alibali, & Goldin-Meadow, 1998).

Indeed, gesture-speech mismatches in a variety of domains have been associated with children on the verge of learning, serving as a signal that the learner is in a state of transition and is particularly receptive to instruction on a task, compared to children who convey the same procedure in speech and gesture (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Goldin-Meadow, Alibali, & Church, 1993; Perry et al., 1988). Thus, gesture-speech mismatches have been labeled as an index of a child’s openness to instruction or “readiness to learn.” These foundational studies on gesture-speech mismatches provided the impetus for research investigating the role that gesture plays in learning, with several researchers arguing that gesture is not merely an index of a child’s “readiness to learn,” but actually plays a key role in changing children’s knowledge (e.g., Alibali & Goldin-Meadow, 1993). However, Cook
and Goldin-Meadow (2006) noted that in most studies, children have not been encouraged to gesture, but instead gestured spontaneously during their explanations of solutions to problems. Therefore, those gestures could simply be a reflection of children’s “readiness to learn” but not play an active role in changing children’s knowledge.

Cook and Goldin-Meadow (2006) posited that gesture can create new knowledge, not just act as an index of children’s “readiness to learn” prior to instruction. In their study, they successfully increased the rate at which children gestured when explaining their solutions to math equivalence problems by exposing children to teachers’ gestures during a lesson on math equivalence. The researchers then examined the relation between children’s gesture production and learning. Children who produced gestures of their own after viewing the teachers’ gestures were more likely than those who did not produce gestures to both retain and generalize the knowledge gained during instruction, as shown in performance on a posttest including familiar and novel problems. These results suggest gesturing can help children benefit from a lesson, particularly gesturing a correct problem-solving strategy as demonstrated by the instructor, above and beyond expressing that correct strategy with speech alone. However, the results of this study depended on children choosing to gesture on their own. Thus, because not all of the children who observed the teachers’ gestures actually produced that same gesture themselves, the gestures children produced could have possibly been a reflection of their “readiness to learn,” rather than a causal factor in the learning process. To address this concern, Broaders, Cook, Mitchell, and Goldin-Meadow
(2007) performed a more direct experimental manipulation of gesture. They found that children who were simply told to gesture during their explanations of their solutions to math equivalence problems were able to add new problem-solving strategies to their repertoires through gesturing and showed a greater benefit from instruction compared to children told not to gesture.

In more recent studies, researchers have investigated the role of gesture in learning by asking children to mimic gestures instantiating particular strategies for solving math equivalence problems prior to instruction. Cook, Mitchell, and Goldin-Meadow (2008) tested the hypothesis that gestures play a role in the creation and retention of knowledge by comparing posttest performance among children told to mimic a gesture instantiating a correct, “equalize” strategy for solving math equivalence problems, children told to mimic speech describing that strategy, and children told to mimic both speech and gesture. Prior to instruction, children mimicked their assigned behavior three times (either in speech, gesture, or both), and children also mimicked that behavior before and after solving problems on their own during instruction on how to solve the problems. Children in all three groups performed similarly on an immediate posttest; however, the mimicked behavior made a difference in how well children retained the knowledge gained from instruction. Children who mimicked an equalize gesture (i.e., moving the L hand from L to R under the L side of the problem, pausing, and then moving the R hand from L to R under the R side of the problem) performed better on a delayed follow-up test than those who did not mimic the gesture. Interestingly, the gains children maintained from posttest to follow-up test were not
reliably different between the gesture and “both” groups, suggesting that there was something about mimicking the gesture per se—with or without speech—that led to more robust learning or consolidation of what was learned during instruction.

The results from Cook and colleagues (2008) suggest that gesture plays a role in conceptual change by making learning last. Goldin-Meadow, Cook, and Mitchell (2009) extended these findings by showing that mimicking specific gestures prior to instruction on math equivalence problems not only helps children maintain a correct, learned strategy, but also helps them generate a correct problem-solving strategy on their own. Children in this study were taught to mimic gestures that instantiated a different correct strategy than the strategy teachers taught in the lesson, in order to examine whether children’s gestures alone prior to instruction can create new ideas. Children were told to mimic a grouping gesture (i.e., forming a “V” with the R hand under the grouped addends on the L side of the problem whose sum is the correct solution, pausing, and then pointing with the R index finger to the blank on the R side of the problem) before and during instruction on the equalize strategy (making the two sides of the equation sum to the same total). Children who were told to mimic this grouping gesture performed better on a posttest than children who were not told to mimic the grouping gesture (Goldin-Meadow et al., 2009).

These studies provide evidence that gesture can play a causal role in conceptual change, though it is unclear how. Literature suggests that gesture provides a representational format that places less demand on working memory than expressing information in speech alone (e.g., Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).
Thus, gesture is thought to free cognitive resources and reduce effort, decreasing errors in the way information is encoded (Ballard, Hayhoe, Pook, & Roe, 1997). Gesture has also been found to produce more robust memory traces than simply speech, perhaps by activating relevant motoric information. For instance, when speakers act out an event described in a sentence with their hands they remember it better than if they simply read that sentence (von Essen & Nilsson, 2003).

Although recent studies on children’s understanding of math equivalence (e.g., Cook et al., 2008; Goldin-Meadow et al., 2009) support the idea that gesturing facilitates learning, the gestures children have mimicked in these studies have all been relational gestures that move children’s attention back and forth across the equal sign. It is, therefore, unclear how the gestures facilitate learning. I posited that the benefits of these gestures could be due to the relational eye movements that accompany the gestures, rather than to the gestures themselves. Indeed, both gestures and eye movements have been identified in the literature as physical embodiments of cognition, engaging the external environment and grounding learning so that the load on working memory is reduced (Ballard et al., 1997).

Many researchers have investigated associations between eye movements and cognition. Early investigations of perception in chess players have found that eye glances correspond to “chunks” of extracted information in spatial working memory, suggesting that these patterns of eye movements may help players access chess moves that are stored in memory (Chase & Simon, 1973). In a recent investigation, Grant and Spivey (2003) showed that participants’ eye movements predict correct problem
solving. In their study, they used animation (visual pulsing) to induce problem solvers to fixate on the critical feature of a problem (as previously revealed in successful participants’ eye movements). Results indicated that drawing problem solvers’ attention in this way can help solvers develop problem-solving insights. Thus, participants’ eye movements may serve as an embodied physical mechanism that stimulates new ways of thinking about a problem (Grant & Spivey, 2003; but see van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009 for an alternative view). Thomas and Lleras (2007) provided additional evidence for the link between eye movements and cognition in a study that manipulated participants’ eye movements. They showed that directing eye movements on a tracking task in a pattern that embodies a correct solution leads to successful problem solving (Thomas & Lleras, 2007). Additionally, research conducted with adults indicates that looking back and forth across the equal sign is correlated with correct strategies to solve math equivalence problems (Chesney, McNeil, Brockmole, & Kelley, 2013). Drawing upon this recent evidence suggesting that guiding attention guides thought, I theorized that the beneficial effects of gesture on learning of math equivalence could be driven, in part, by eye movements that embody relational thinking.

The present study was designed to directly compare the effects of mimicking gestures to the effects of mimicking the eye movements that accompany those gestures. It built off of Cook et al.’s (2008) design by using both “speech only” and “gesture” conditions and comparing them to an eye movement condition. I hypothesized that children in the eye movement condition would perform better than
children in the speech only condition and similarly to children in the gesture condition, thus demonstrating that the beneficial effects of gesture are not due to the hand movements themselves, but rather to a more general mechanism that co-occurs with gesture. I also hypothesized that the number of times children’s eyes moved back and forth across the equal sign (coded from video) would be an important predictor of learning.
CHAPTER 2:  
METHOD

2.1 Participants

Participants were 107 children (ages 7 yrs 7 mos – 9 yrs 10 mos) from local schools and after-school programs in the Midwestern United States, recruited through letters and permission slips sent to parents; no children were excluded. 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). Because the goal of this study was to examine how different instructional conditions affect children’s learning from a lesson on math equivalence, the final sample for analysis was limited to children who solved all pretest problems incorrectly (69% of the total sample). Additionally, two of those children did not return for the second session, and two other children were excluded from the sample because they could not complete all of the tasks. The final sample of participants included 70 children (34 boys, 36 girls; \( M \) age = 8 years, 9 months). The race/ethnicity of the sample was 17% African-American or black, 4% Asian, 7% Hispanic or Latino, 6% Other, and 66% white.

Sessions were conducted in a quiet room in a research lab, a local school, and a local afterschool program. 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 may want to end the session.

2.2 Design

The study was a pretest-intervention-posttest design, with a 4-week follow-up, akin to the design used by Cook and colleagues (2008). Each child participated in two sessions; the first session lasted approximately 45 minutes, whereas the second session lasted approximately 30 minutes. The first session consisted of four phases: pretest, pre-instruction, instruction, and posttest. The second session consisted of a follow-up test and a brief lesson on math equivalence tailored specifically to the child’s needs. Both sessions were videotaped, so that we could study the strategies children used when solving the problems.

2.3 Experimental Conditions

Children were randomly assigned to one of three conditions: speech only ($n = 23$), gesture ($n = 24$), or eye movement ($n = 23$). Each child received the same instruction on math equivalence and the same assessments; the only aspect that varied was the behavior children were asked to mimic during the pre-instruction and instruction phases. One experimenter served as the lesson facilitator during the pre-instruction and instruction phases, and a different experimenter, who was blind to the
child’s condition, served as the tester, administering the pretest, posttest, and follow-up test. The lesson facilitator also taught the brief lesson at the end of the second session.

2.3.1 Speech Only

In the speech only condition, children were shown a video of a teacher standing in front of a problem saying the phrase “I want to make one side equal to the other side.”

2.3.2 Gesture

In the gesture condition, children were shown a video of a teacher standing in front of a problem saying the same phrase while simultaneously producing a relational, equalize gesture (moving the L hand from L to R under the L side of the problem, pausing, and then moving the R hand from L to R under the R side of the problem) (See Figure 1 for screenshots from the video of the teacher demonstrating this behavior).

2.3.3 Eye Movement

In the eye movement condition, children were shown a video of a teacher standing in front of a problem saying the same phrase while simultaneously moving their eyes across the problem in a way that simulated the eye movements that would co-occur with gesture in the gesture condition. To encourage eye movements, an arrow moved underneath each side of the problem during the video (from L to R under the L side of the problem, disappearing briefly, and then from L to R under the R side of the problem) (See Figure 2 for screenshots from the video of the teacher demonstrating this behavior).
behavior). Thus, this condition emulated all aspects of the gesture condition except the actual hand gestures.

2.4 Procedure

2.4.1 Pretest

Children solved four math equivalence problems with equivalent addends on each side of the equal sign (5 + 4 + 6 = __ + 6, 3 + 5 + 9 = __ + 9, 8 + 4 + 3 = 8 + __, 7 + 5 + 8 = 7 + __). An experimenter presented each equation at a whiteboard and said, “Try your best to solve the problem, and then write the number that goes in the blank.” After children wrote a number in the blank for each problem, the experimenter 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).

2.4.2 Pre-instruction

The lesson facilitator showed children a video of a teacher demonstrating a behavior and asked the children to mimic that behavior for three problems of the format $a + b + c = __ + c$. The behavior a given child was asked to mimic depended on which condition he or she was in (as described above; see also Figures 1 and 2). The facilitator said “Now we’re going to practice saying some words [and moving our hands/eyes] that will help you think about the problems you were solving earlier. We will watch a video of a teacher standing in front of a problem. She will say words [and move her hands/eyes in ways] that help her think about the problem. I want you to
watch the teacher and then see if you can copy what she does, okay?” In the eye
movement condition, children were also told “An arrow will appear on the screen to
help her know how to move her eyes” during the introduction to the video to help
encourage learning of the eye movements. In all conditions, the experimenter showed
the video of the teacher performing the behavior twice to ensure that all children
understood the procedure. After each video the children were asked “Now can you say
the words [and do the hand/eye movements] just like the teacher did?” Children were
shown the videos on a laptop, and then children were presented with two additional
math equivalence problems alone on the laptop screen for them to practice doing the
behavior on their own. For those two additional problems, children were asked “Can
you say the words [and do the hand/eye movements] to this problem?” After each time
children practiced the behavior in pre-instruction, they were encouraged to remember
that behavior because they would be asked to do it again later on.

Videos of children’s faces and computer screens in all conditions were recorded
during the pre-instruction and instruction using customized software and the laptop
computer’s built-in camera. The two videos were automatically synchronized so that the
children’s faces (including eye movements) and what they were seeing on the laptop
(e.g., a video of the teacher, a particular problem, etc.) were temporally aligned. Hence,
eye movements (coded from videos) could be connected to every point in the pre-
instruction and instruction phases of the session. In this way, I could not only ensure
that children in the eye movement condition were doing the eye movement behavior as
encouraged, but I could also analyze the role of children’s eye movements in problem solving across all conditions.

2.4.3 Instruction

Even though conditions differed in terms of what they saw during the pre-instruction phase, all children saw the same instruction (cf. Cook et al., 2008). Children watched a video of a teacher explaining how to use the equalize strategy to solve six more problems of the same type ($a + b + c = \_ + c$, cf. Goldin-Meadow et al., 2009). For each of these problems, the teacher described the strategy “I want to make one side equal to the other side” both before and after solving the problem. Each time the teacher said that phrase, an arrow moved underneath each side of the problem while the teacher made the relational, “equalize” gesture (moving the L hand from L to R under the L side of the problem, pausing, and then moving the R hand from L to R under the R side of the problem) (See Figure 3 for screenshots from the video of an instruction example). Thus, children in all conditions were exposed to the equalize strategy 12 times in speech, gesture, and eye movements (encouraged by the arrows). This instruction ensured that all children were exposed to the same representations of equivalence and increased the chance that at least some children would learn how to solve math equivalence problems of that type.

After each of the teacher videos, children saw another problem presented alone on the laptop screen. First, children were asked to reproduce the behavior they practiced during pre-instruction. Next, the lesson facilitator placed a transparency sheet over the laptop screen and asked children to solve the problem using a transparency
marker. Children were not given any feedback about correctness. After taking the marker away, the lesson facilitator asked children to reproduce the behavior they practiced during pre-instruction again. Children who produced behaviors other than what they had practiced during pre-instruction were reminded to only produce the behaviors they were instructed to mimic.

2.4.4 Posttest

Immediately after the instruction phase, children completed a posttest administered by the tester that included the pretest equations (see above) along with a larger assessment of understanding of math equivalence that included transfer equations, equation encoding, defining the equal sign, and identifying the two sides of a math equivalence problem. These additional measures were included to determine the depth of children’s learning across conditions.

Transfer equations included math equivalence problems that differed in terms of surface features (7 + 4 + 6 = __ + 3, 6 + 2 + 8 = 5 + __, 1 + 5 = __ + 2, 6 − 1 = 3 + __). Children’s problem-solving strategies on all equations were coded as correct or incorrect based on a system used in previous research (e.g., McNeil & Alibali, 2004; Perry et al., 1988). For most problems, correctness could be inferred from the solution itself (e.g., for the problem 7 + 5 + 8 = 7 + __, a solution of 27 indicated an incorrect “add all” strategy and a solution of 13 indicated a correct strategy). If the solution was ambiguous, then strategy correctness was coded based on children’s verbal explanation (e.g., for the problem 7 + 5 + 8 = 7 + __, the explanation “I added 7 plus 5” indicated an incorrect strategy and the explanation “I added 5 plus 8” indicated a correct strategy).
To assess equation encoding, children were asked to reconstruct four equations with operations on both sides of the equal sign (e.g., $2 + 3 + 6 = 2 + ___$) using paper and pencil after viewing each for five seconds (cf. Chase & Simon, 1973; Siegler, 1976). 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. Correctness was coded based on a system used in previous research (e.g., McNeil & Alibali, 2004; Rittle-Johnson & Alibali, 1999). Common errors included converting the problem to a traditional arithmetic problem in the “operations equals answer” format (e.g., reconstructing $4 + 5 = 3 + ___$ as “$4 + 5 + 3 = ___$”), omitting the plus sign on the right side of the equal sign (e.g., reconstructing $4 + 5 = 3 + ___$ as “$4 + 5 = 3 ___$”), and omitting the equal sign (e.g., reconstructing $7 + 1 = ___ + 6$ as “$7 + 1 ___ + 6$”).

To assess defining the equal sign, children responded to a set of interview questions about the equal sign. The experimenter pointed to an equal sign presented alone on a piece of paper and asked: (1) “What is the name of this math symbol?” (2) “What does this math symbol mean?” and (3) “Can it mean anything else?” (cf. Baroody & Ginsburg, 1983; Behr et al., 1980; Knuth et al., 2006). Responses were categorized according to a system used in previous research (e.g., Knuth et al., 2006; McNeil & Alibali, 2005a, b). We were interested in whether or not children defined the equal sign relationally as a symbol of math equivalence (e.g., “Two amounts are the same”). After children responded to the interview questions, they were asked to rate the “smartness” of eight fictitious children’s responses on a continuous scale ranging from “not so smart” to “very smart.” Of the eight fictitious children’s responses, two represented a
“sameness” relational understanding (“something is equal to another thing” and “two amounts are the same”), two represented a “substitutive” relational understanding (“the things on the right side of it can be swapped with the things on the left side of it” and “the two sides of a math problem can be exchanged”), two represented an operational understanding (“the answer to the problem” and “the total”), and two were distractors (“the end of the problem” and “repeat the numbers”) (cf. Jones, Inglis, Gilmore, & Dowens, 2012; Rittle-Johnson & Alibali, 1999).

To assess identifying the two sides of a math equivalence problem, children were shown an equation in the same format as the equations seen during instruction (\(a + b + c = \_ + c\)). The experimenter then asked “There are two sides to this problem. Can you tell me what you think might make up each side? So, what numbers and math symbols make up one side and what numbers and math symbols make up the other side?” (cf. Rittle-Johnson & Alibali, 1999). Children then responded to the same questions for a second equation of the form \(a + b + c = a + \_\).

2.4.5 Follow-up test

Approximately four weeks after the first session, children completed a follow-up test identical to the posttest.
CHAPTER 3:
RESULTS

3.1 Effects of Condition

3.1.1 Posttest Performance

Recall that none of the children in this sample solved equations correctly on the pretest. The learning rate was high, with most children (73%) solving at least one posttest equation correctly (out of 8) and a majority (60%) solving at least one of the first four posttest equations correctly. This learning rate was comparable to the learning rate in Cook et al.’s (2008) study, 52%, in which students received live instruction. As in Cook et al. (2008), there was no evidence of significant differences in performance solving equations among conditions during instruction, $F(2, 67) = 0.30, p = .74$, or on the immediate posttest, $F(2, 67) = 0.02, p = .98$.

3.1.2 Follow-up test Performance

Following Cook et al. (2008), I tested if children across conditions differed in how well they maintained the knowledge gained during instruction over the 4-week delay. I conducted an ANCOVA with condition (speech only, gesture, or eye movement) as the independent variable, number of posttest equations correct (out of 8) as the covariate, and number of follow-up equations correct (out of 8) as the dependent variable. Not
surprisingly, posttest performance significantly predicted follow-up performance, \( F(1, 66) = 52.93, p < .001, \eta^2_p = .45 \), with higher posttest equation solving performance associated with higher follow-up equation solving performance. The effect of condition was also significant, \( F(2, 66) = 3.90, p = .025, \eta^2_p = .11 \). Contrary to my hypothesis, it was children in the gesture condition who did not retain the knowledge they had gained during instruction (see Figure 4). Simple contrasts indicated that children in the gesture condition had significantly lower retention scores than children in both the eye movement condition, \( p = .035, \text{Cohen's } d = .63 \), and the speech only condition, \( p = .011, \text{Cohen's } d = .76 \). Children in the eye movement condition did not differ from children in the speech only condition, \( p = .66 \).

Findings were robust, even when I varied aspects of the analysis. For example, conclusions were unchanged (the effect of condition remained at least marginally significant) when I limited the analysis to the four equations that matched the pretest equations, \( F(2, 66) = 2.95, p = .059, \eta^2_p = .08 \), when I limited the analysis to only the four transfer equations, \( F(2, 66) = 4.35, p = .017, \eta^2_p = .12 \), and when I excluded the children who did not show evidence of learning from instruction (i.e., children who did not solve at least one posttest equation correctly), \( F(2, 47) = 8.23, p = .001, \eta^2_p = .26 \). Conclusions also remained unchanged in each of those analyses when I excluded the children who demonstrated a correct strategy in gesture at the pretest. The effect of condition was significant for the original ANCOVA predicting follow-up equation-solving performance (out of 8), \( F(2, 53) = 5.56, p = .006, \eta^2_p = .17 \), when I limited the analysis to the four equations that matched the pretest equations, \( F(2, 53) = 4.07, p = .023, \eta^2_p = .13 \), when I
limited the analysis to only the four transfer equations, $F(2, 53) = 6.04, p = .004, \eta_p^2 = .19$, and also when I excluded the children who did not show evidence of learning from instruction (i.e., children who did not solve at least one posttest equation correctly), $F(2, 39) = 7.07, p = .002, \eta_p^2 = .27$.

There were no other significant differences across conditions in equation encoding, equal sign definition, or identifying the two sides of an equation at either posttest or follow-up test. Additionally, comparing performance across conditions on a composite measure of understanding equivalence did not yield any significant differences. Table 1 displays means and standard deviations for performance on each of the measures of understanding of mathematical equivalence during instruction, posttest, and follow-up by condition.

3.1.3 Level of Adherence to Condition

To further probe these unexpected effects of the gesture condition, I coded children’s level of adherence to the relational equalize gestures that were modeled by the teacher in the pre-instruction and instruction phases. Equalize gestures were coded using a system that has been established in previous work on children’s gestures to math equivalence problems (cf. Alibali & Goldin-Meadow, 1993). Children were given a score of “1” on each equation if they ever completed the full equalize gesture with two different hands as demonstrated, “0.5” if they made a different equalize gesture (equalize gestures are gestures that distinguish the two sides of the equation, for example changing hand shape or putting the hand down in between the left and right sides of the equation), and “0” if their gesture was not an equalize gesture. Children’s
scores across all nine problems were added together for a total level of adherence score. Children’s adherence in the gesture condition was far from perfect, though highly variable, with a mean level of adherence of 4.88 (SD = 3.16). Fifty percent of children distinguished the sides in their gesture on all problems, and half of those children completed the full gesture with two hands as demonstrated. Most children did make some sort of equalize gesture, however. Fifty-four percent of children made an equalize gesture on at least half of the problems, and only 8% of children never made an equalize gesture. However, I did not find evidence that the degree of adherence was associated with retention. With the sample of children in the gesture condition, I conducted a multiple regression with level of adherence (out of 9) as the independent variable, number of posttest equations correct (out of 8) as the covariate, and number of follow-up equations correct (out of 8) as the dependent variable. Level of adherence was not a significant predictor of follow-up equation solving performance, after controlling for posttest performance, \( b = .22, t(23) = 1.08, p = .29 \). Thus, there was no evidence that children’s level of adherence to the modeled equalize gesture affected retention of learning.

In order to test a full analysis of adherence to condition as a predictor of children’s performance, adherence in the speech and eye movement conditions was coded in a similar manner as the gesture condition for each equation seen in the pre-instruction and instruction phases (see above). For the speech condition, children were given a score of “1” on each equation if their phrase ever indicated one side “equal to” another side, “0.5” if they used “equal,” “equals,” or “equal as” in the phrase instead of
“equal to,” and “0” if they did not indicate one side and another (e.g., “make an equal side”). For the eye movement condition, children were given a score of “1” on each equation if they ever directly switched between looking to one side and looking to the other side, “0.5” if they switched directions while looking above the problem, and “0” if they did not switch directions. As done for the gesture condition, children’s scores across all nine problems were added together for a total level of adherence score. One child in the eye movement condition was excluded from this analysis because their eye movements were not properly recorded from the laptop camera and could not be coded, thus, the final sample for this analysis included 69 children.

For this analysis, I first conducted an ANOVA with condition (speech only, gesture, or eye movement) as the independent variable and level of adherence (out of 9) as the dependent variable. The effect of condition was significant, $F(2, 66) = 16.26, p < .001, \eta^2_p = .33$. Simple contrasts indicated that children in the gesture condition had significantly lower level of adherence scores ($M = 4.88, SD = 3.16$) than children in both the speech only condition, $p < .001 (M = 7.91, SD = 1.79)$, and the eye movement condition, $p < .001 (M = 8.34, SD = 1.36)$. Children in the speech only condition did not differ from children in the eye movement condition, $p = .84$. This significant effect of condition held and the same pattern of results was seen when I collapsed my three-level coding scheme into a two-level one, in which any behavior that qualified as “.5” was given “1” (see above descriptions of the coding scheme for level of adherence in each condition), $F(2, 66) = 7.86, p = .001, \eta^2_p = .19$. 
I then conducted a multiple regression analysis with level of adherence (out of 9) as the independent variable, condition (speech only, gesture, or eye movement) as the covariate, and number of posttest equations correct (out of 8) as the dependent variable. After controlling for condition, level of adherence was a significant predictor of posttest equation solving performance, $b = .38$, $t(65) = 2.17$, $p = .033$. Thus, level of adherence to the exact behavior predicted immediate learning.

A final step-wise regression analysis was conducted with number of follow-up equations correct (out of 8) as the dependent variable. In the first step, condition (speech only, gesture, or eye movement) and number of posttest equations correct (out of 8) were entered as the covariates. In the second step, level of adherence (out of 9) was entered as an independent variable. In the third step, an interaction between condition and level of adherence was entered as another independent variable to test for a moderation effect. The $R^2$ change for the second model including level of adherence ($R^2 \Delta = .003$) was not significant, $p = .52$, and the $R^2$ change for the third model including the level of adherence by condition interaction ($R^2 \Delta = .011$) was also not significant, $p = .52$. Thus, although level of adherence was a significant predictor of immediate learning (posttest performance), level of adherence was not a significant predictor of maintenance of learning in the long-term, and neither was the level of adherence by condition interaction.
3.2 Eye Movement Analyses

3.2.1 Eye Movements Across Conditions

Recall that the primary rationale for conducting the present study was related to participants’ eye movements. I hypothesized that participants’ eye movements back and forth across the equal sign during pre-instruction and instruction would be an important predictor of learning. Thus, I first examined if condition was a significant predictor of children’s relational eye movements during pre-instruction and instruction. Children’s eye movements were recorded from the laptop and coded in three types of situations: (1) during self-production of the behavior in both the pre-instruction and instruction phases, (2) while solving problems during instruction, and (3) while watching videos of the teacher solving problems and doing the behavior during instruction. For each situation, I calculated the number of times children looked back and forth across the equal sign on each problem (i.e., related the two sides of the problem), and then averaged across all problems seen during pre-instruction and instruction. Additionally, eye movements during video instruction were unable to be coded for three participants because of recording malfunctions.

First, I performed an ANOVA with average number of relational eye movements during children’s self-production of the behavior as the dependent variable and condition (speech only, gesture, or eye movement) as the between-subjects factor. There was a significant effect of condition, $F(2, 66) = 32.88, p < .001, \eta_p^2 = .50$. Simple contrasts indicated that children in the speech only condition made significantly fewer relational eye movements in the self-production of their behavior on average ($M = 0.38$,
than did children in the eye movement condition, \( p < .001 \) (\( M = 0.96, SD = 0.37 \)), and the gesture condition, \( p < .001 \) (\( M = 0.95, SD = 0.21 \)). Children in the eye movement condition did not differ from children in the gesture condition, \( p = .89 \). This result was expected based on the characteristics of the behavior to be mimicked and my hypothesis concerning the eye movements that accompany gesture. There were no significant differences among conditions for either relational eye movements while solving problems during instruction, \( F(2, 66) = 0.62, p = .54, \eta_p^2 = .02 \), or relational eye movements while watching videos of the teacher, \( F(2, 63) = 1.44, p = .25, \eta_p^2 = .04 \).

3.2.2 Eye Movements and Performance

Next, I considered whether children’s relational eye movements during pre-instruction and instruction predicted their performance, ignoring condition. In a regression analysis, I found that the average number of eye movements back and forth across the equal sign while solving problems during instruction was a marginally significant predictor of performance on the instruction problems, \( b = .40, t(67) = 1.96, p = .054 \).

As hypothesized, relational eye movements were positively associated with use of correct strategies to solve problems (as seen in the regression predicting correct performance). Building upon this result, I was interested in whether relational eye movements were also negatively associated with use of typical arithmetic strategies for problems solved incorrectly. Because children were not asked to explain how they got their solutions for each instruction problem, strategies used during instruction were coded based only on the number written in the blank. For example, for the problem 4 +
5 + 7 = __ + 7, “16” was coded as “Add to Equal Sign” and “19” was coded as “Add All” (both typical arithmetic strategies used most often by children). Some other incorrect strategies seen for the problem above, for example, were “12” (summing the wrong two numbers) or “4” (just carrying over a number from the left side). Solutions were coded as reflecting a particular strategy if they were within ±1 of the solution that would be achieved with that particular strategy. In order to determine whether relational eye movements were negatively associated with use of typical arithmetic strategies (“Add to Equal” or “Add All”) for problems solved incorrectly, I used a mixed-effects generalized linear model with a logit link function (i.e., a repeated measures logistic regression model) using the lme4 (Bates, Maechler, & Bolker, 2011) package in the R language and environment for statistical computing (R Core Team, 2014). The dichotomous outcome of typical arithmetic strategy used (yes/no) was modeled as the log odds. Examining only incorrect trials, I first constructed a model with only an intercept, both as a fixed effect (the average across individuals) and a random effect (the individual-specific deviation from the fixed effect). I then constructed a second model including this intercept along with relational eye movements during solving. In this second model, the number of relational eye movements on a particular problem negatively predicted (though it was only marginally significant) use of a typical arithmetic strategy, \( b = -0.48, z = -1.89, p = .058 \). Model comparisons indicated that this model fit the data significantly better than the first model, \( p = .016 \), which did not include relational eye movements during solving. It seems that children’s eye movements back and forth between sides of the problem (i.e., relating the sides) while solving instruction problems were associated
not only with use of correct strategies to solve the problems, but also with incorrect strategies other than those seen in typical arithmetic (i.e., just adding up numbers before the equal sign or adding all of the numbers present), a potential indication of transition towards correct understanding.

I also examined eye movements during children’s own production of their behavior during pre-instruction and instruction as an independent variable. In a regression analysis, I found that the average number of relational eye movements during children’s self-production of the behavior was a marginally significant predictor of posttest performance identifying the sides of an equation of the same format as presented during instruction, \( b = .25, t(67) = 1.72, p = .09 \). Within the sample of learners (children who solved at least one posttest problem correctly), the average number of relational eye movements during self-production of the behavior was a significant predictor of posttest performance identifying the sides of an equation of the same format as presented during instruction, \( b = .37, t(48) = 2.07, p = .044 \).

3.3 Post-Hoc Analyses

Recall that children in the gesture condition had significantly lower scores on the level of adherence scale than did children in the other conditions. This lower level of adherence may have been because the particular gesture towards a computer may have been unnatural or awkward for children. Children in the present study found it difficult to mimic the exact gesture to the laptop screen, and some even voiced these difficulties (e.g. “This is hard.”). As a measure of children’s difficulty mimicking the exact behavior
in each condition, I calculated the number of times children looked up at the lesson facilitator during their production of the behavior summed across each instance of the behavior. I then tested if children differed in the number of times they looked up at the facilitator during their production of the behavior across conditions. I conducted an ANOVA with condition (speech only, gesture, or eye movement) as the between-subjects factor and number of looks towards the lesson facilitator during production of their behavior as the dependent variable. I found a significant effect of condition, $F(2, 66) = 5.59, p = .006, \eta_p^2 = .15$. Simple contrasts indicated that children in the gesture condition had significantly more looks towards the lesson facilitator during their production of the behavior ($M = 2.42, SD = 0.45$) than children in both the eye movement condition, $p = .004 (M = 0.46, SD = 0.47)$, Cohen’s $d = .81$, and the speech only condition, $p = .008 (M = 0.65, SD = 0.46)$, Cohen’s $d = .71$. Children in the eye movement condition did not differ from children in the speech only condition, $p = .77$.

The greater number of looks towards the lesson facilitator in the gesture condition than in either of the other two conditions as well as the lower mean level of adherence to the exact behavior may indicate that the particular gesture may have felt unnatural to children, or that children did not understand what they were supposed to do. Attempts to produce gestures that feel unnatural in instructional settings may increase cognitive load and may have a disproportionately negative effect on children with low expertise in a content area (like those in the present study) (e.g., Post, van Gog, Paas, & Zwaan, 2013).
To test the possibility that the unnatural gesture towards a computer screen added extraneous cognitive load to the learning task, I calculated the time it took for children to solve each problem during instruction for all conditions as a proxy measure of cognitive load (cf. Sweller, van Merrienboer, & Paas, 1998). First, I took the average time for all problems solved, and conducted an ANOVA with average solution time as the dependent variable and condition (speech only, gesture, or eye movement) as the between-subjects factor, and the effect of condition was not significant, $F(2, 67) = 0.37$, $p = .69$, $\eta^2_p = .01$. Within the sample of learners (children who solved at least one posttest problem correctly), the effect of condition was also not significant, $F(2, 48) = 0.80$, $p = .46$, $\eta^2_p = .03$, though the gesture condition had a relatively higher mean solution time (14.59 seconds, $n = 16$) than did the speech only condition (11.97 seconds, $n = 17$) and the eye movement condition (11.64 seconds, $n = 18$). Finally, I calculated a separate variable averaging solution time for only the problems that were solved correctly for the sample of learners. The effect of condition in this ANOVA was not significant, $F(2, 40) = 1.51$, $p = .23$, $\eta^2_p = .07$, however, the gesture condition still had a relatively higher mean time spent solving problems correctly in this small sample (15.46 seconds, $n = 13$) than did the speech only condition (11.41 seconds, $n = 15$) and the eye movement condition (10.84 seconds, $n = 15$).

To test whether the gesture manipulation was differentially effective depending on children’s initial level of entrenchment, I calculated the number of different strategies children used during the pretest. It could have been that the gesture manipulation was more effective for children who used a greater variety of strategies on
pretest (and thus, were less entrenched prior to instruction). First, I conducted an ANOVA with number of strategies used at pretest as the dependent variable and condition (speech only, gesture, or eye movement) as the between-subjects factor in order to confirm that there were no preexisting differences among the conditions, $F(2, 67) = 2.16, p = .12, \eta_p^2 = .06$. Next, limiting the sample to only children in the gesture condition ($n = 24$), I conducted a regression analysis with number of follow-up equations correct (out of 8) as the dependent variable, number of posttest equations correct (out of 8) as the covariate, and number of strategies used at pretest as the independent variable. Number of strategies used at pretest was not a significant predictor of retention for children in the gesture condition, $b = -.85$, $t(21) = -.90$, $p = .38$.

In a related analysis, I coded children’s gestures during their explanations of solutions on the pretest problems using a system that has been established in previous work on children’s gestures to math equivalence problems (cf. Alibali & Goldin-Meadow, 1993). It could have been that the gesture manipulation was more effective for children who conveyed different strategies in gesture and speech on the pretest (i.e., produced a gesture-speech mismatch). With the sample of children in the gesture condition, ($n = 24$), I conducted an ANCOVA with number of follow-up equations correct (out of 8) as the dependent variable, number of posttest equations correct (out of 8) as the within-subjects factor, and whether children produced a gesture-speech mismatch at pretest as the between-subjects factor. There was not a significant effect of gesture-speech mismatch at pretest on retention, $F(1, 21) = 0.01$, $p = .93$. 

31
I hypothesized that children in the eye movement and gesture conditions would learn and retain more from instruction on math equivalence when compared to children in the speech only condition. However, contrary to my expectations, children in the gesture condition actually retained less of the knowledge they had gained during instruction when compared to children in the other two conditions. Overall, these results suggest that there may be some limits to the benefits of gesture during instruction. At the same time, however, results provided some support for the hypothesis that children’s relational eye movements back and forth across the equal sign during pre-instruction and instruction are associated with learning. Specifically, children who produced more relational eye movements during the pre-instruction and instruction phases, on average, solved more problems correctly during instruction.

Previous research has detailed the benefits of gesture on learning in various contexts – the benefits of spontaneous gesture during a lesson (Cook & Goldin-Meadow, 2006), the benefits of being told to gesture (Broaders et al., 2007), and the benefits of self-producing gestures of correct strategies (Cook et al., 2008, Goldin-
Meadow et al., 2009). Thus, it is important to determine what it was about the present study that reversed these benefits.

There are several potential reasons why the expected benefits of gesture (on both learning and retention) were not found in the present study, and each provides fodder for future research.

First, the physical presence of the teacher may moderate the effects of gesture on learning and retention. Perhaps watching a video of a teacher gesturing and mimicking that gesture may be experienced differently than watching a teacher gesture in real life and mimicking that gesture. It could have felt uncomfortable for children to mimic the gestures of someone who was not there to see them. Although work by Cook, Duffy, and Fenn (2013) has shown positive learning effects from watching gestures on a videotaped lesson on math equivalence, the lack of the physical presence of the teacher in this study may be impacting the depth of the instantiation for children’s own mimicry of gestures. There is a learning phenomenon called “social gating” in which learners are especially attuned to information presented in the context of social interaction, and, thus, knowledge acquisition is enhanced (Sage & Baldwin, 2009). Children in Cook et al.’s (2008) study may have been more engaged and motivated in the social interaction of taking turns with a live teacher and jointly attending to the problem, and may have been more deeply focusing and encoding the strategy as they were performing the gesture during their “turn.” Indeed, research has documented the importance of the visibility of the “listener” in gestural communication (e.g., Alibali, Heath, & Myers, 2001). Gesturing to a communicative partner can be experienced differently than
gesturing when a partner is not present, or even not visible (e.g., blocked by a screen). Alibali and colleagues (2001) found that speakers produced representational gestures (gestures that convey semantic content) much more frequently when a partner was present with them than when their partner could not see their gestures. It is possible that without a live teacher present in this study, the gestures children produced did not hold as much meaningful information.

Second, redundancy of information during the instruction video may moderate the effects of gesture on learning and retention. Recall that in the present study, aspects of each condition were included in the instruction (i.e., an arrow moved under the problem while the teacher spoke and gestured). Cook et al.’s (2008) study did not include an eye movement condition, so the instruction only included speech and gesture, without an arrow. Perhaps the arrow that appeared underneath the problem attracted attention away from the teacher’s gestures (which were subsequently lower on the screen, below the arrow; see Figure 3) and interfered with children’s mimicry of the behavior in the gesture condition. Redundancy during the instruction video may have resulted in a weaker instantiation of the physical gesture in memory for children in the gesture condition, and thus, a weaker connection to and embodiment of the strategy when performing the behavior themselves. In the literature, a “redundancy effect” occurs when different sources of information are intelligible in isolation, and each source is providing similar information in a different form. The very existence of multiple sources of information takes attention and cognitive resources away from knowledge acquisition (Kalyuga, Chandler, & Sweller, 1999). Beyond the detriment to
retention observed in the gesture condition, the arrow during the instruction video may have provided additional benefit for children in the speech condition. Children in the gesture condition in this study solved about the same number of problems correctly on average during instruction (2.9 out of 6) as did children in the same condition (gesture and speech) in Cook et al.’s (2008) study (2.8 out of 6). However, children in the speech only condition in this study solved a greater number of problems correctly on average during instruction (3.4 out of 6) than did children in the speech condition in Cook et al.’s (2008) study (1.8 out of 6). These performance differences indicate that children in the speech condition in this study were gaining additional knowledge from the teachers’ instruction, perhaps as a result of the extra cue directing attention that the arrow provided. Thus, the arrow during instruction may have been helpful for children in the speech condition but redundant for children in the gesture condition.

Third, the space in which children learn and produce gestures may moderate the effects of gesture on learning and retention. In the present study, children gestured to a problem presented on a laptop after watching a video of a teacher gesturing. In Cook et al.’s (2008) study, the teacher’s gestures, the children’s gestures, and the problems children solved during instruction were all in the exact same space (at a board). Perhaps as the children in Cook et al.’s (2008) study watched the teacher do the gesture during instruction in the exact space at the board that they were about to do it, the children anticipated being imitated by the teacher again after their turn, which may have resulted in deeper encoding of the strategy in memory. Indeed, anticipation of being imitated has been found to facilitate motor movements (Pfister, Dignath, Hommel, &
Kunde, 2013). Also, children in the present study were sitting down in front of a laptop and making fairly small hand movements compared to the Cook et al. (2008) study in which children were standing and making larger gestures. Perhaps larger arm movements are needed for deep embodiment of the equalize strategy.

A final, related point is that the act of gesturing towards a laptop screen may have been unnatural. Asking children in the other conditions to say words and do eye movements towards a laptop screen may have been a more familiar and natural interaction with a laptop screen than the gesture that children were asked to mimic. The fact that gesturing towards a laptop screen was unnatural or awkward for children may have added extraneous cognitive load to the learning task.

Extraneous cognitive load makes processing information during learning more difficult. Students who are burdened by this extraneous load during learning are not able to construct the depth of knowledge that other students may be able to because they cannot devote all of their cognitive resources to the learning process (Sweller et al., 1998). Recall that children in the gesture condition had significantly lower scores on the level of adherence scale than did children in the other conditions, and this unnatural gesture contributing to increased cognitive load may have had a negative effect on children in this study, akin to negative effects found in other studies of children with low expertise in a content area (e.g., Post et al., 2013). Although there were no significant differences across conditions in terms of solution time (a proxy measure of cognitive load, cf. Sweller et al., 1998), the gesture condition did have a higher mean solution time within the smaller sample of learners only. Perhaps with a larger sample of learners
I may have more evidence to support the hypothesis that in this situation the gesture added extraneous cognitive load. In the future, another measure of cognitive load could be a short questionnaire students in all conditions fill out after the session, so that some level of their feelings of awkwardness could be assessed from self-report.

Overall, results provide important data regarding potential limits to the benefits of gesture during instruction. These findings not only advance theory and provide future avenues of study, but also provide educators with an important caveat when designing lessons and learning materials for teaching children the concept of mathematical equivalence.
Figure 1. Gesture Condition Pre-instruction. The teacher moves her left hand from left to right under the left side of the equation (top two pictures), pauses, and moves her right hand from left to right under the right side of the equation (bottom two pictures).
Figure 2. Eye Movement Pre-instruction. The teacher looks while the arrow moves from left to right under the left side of the equation (top two pictures), the arrow disappears briefly, and the teacher looks while the arrow moves from left to right under the right side of the equation (bottom two pictures).
Figure 3. Instruction Example. The equalize strategy is instantiated in speech, gesture, and eye movements (encouraged by the arrows) simultaneously.
Figure 4. Equation Solving Performance at Posttest and Follow-up test by Condition. Error bars represent standard errors.
TABLE 1.

PERFORMANCE ON EACH OF THE MEASURES OF UNDERSTANDING OF MATHEMATICAL EQUIVALENCE DURING INSTRUCTION, POSTTEST, AND FOLLOW-UP BY CONDITION

<table>
<thead>
<tr>
<th>Task &amp; performance measure</th>
<th>Speech Only</th>
<th>Gesture</th>
<th>Eye Movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction equation solving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD) out of 6</td>
<td>3.43 (2.61)</td>
<td>2.88 (2.72)</td>
<td>3.00 (2.43)</td>
</tr>
<tr>
<td>Posttest equation solving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD) out of 8</td>
<td>3.70 (3.11)</td>
<td>3.79 (3.75)</td>
<td>3.61 (3.12)</td>
</tr>
<tr>
<td>Posttest equation encoding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD) out of 4</td>
<td>1.61 (1.12)</td>
<td>1.17 (1.17)</td>
<td>1.13 (1.06)</td>
</tr>
<tr>
<td>Posttest equal sign definition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% who defined relationally</td>
<td>34.8</td>
<td>25.0</td>
<td>22.7</td>
</tr>
<tr>
<td>Posttest identifying the sides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% who distinguished sides</td>
<td>30.4</td>
<td>54.2</td>
<td>30.4</td>
</tr>
<tr>
<td>Follow-up equation solving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD) out of 8</td>
<td>4.09 (3.66)</td>
<td>2.21 (3.16)</td>
<td>3.70 (3.40)</td>
</tr>
<tr>
<td>Follow-up equation encoding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (SD) out of 4</td>
<td>1.61 (1.34)</td>
<td>1.46 (1.10)</td>
<td>1.57 (1.24)</td>
</tr>
<tr>
<td>Follow-up equal sign definition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% who defined relationally</td>
<td>21.7</td>
<td>25.0</td>
<td>13.0</td>
</tr>
<tr>
<td>Follow-up identifying the sides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% who distinguished sides</td>
<td>34.8</td>
<td>33.3</td>
<td>26.1</td>
</tr>
</tbody>
</table>
REFERENCES


