© 2016 Neet Priya Bajwa

DEVELOPING A RELATIONAL MEANING OF THE EQUAL SIGN: EFFECTS OF USING A BALANCE ANALOGY IN A GAME-BASED VIRTUAL ENVIRONMENT

BY

NEET PRIYA BAJWA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Curriculum and Instruction
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Sarah Lubienski, Chair
Professor Michelle Perry, Director of Research
Assistant Professor Robb Lindgren
Assistant Professor Daniel Hoffman

ABSTRACT

Understanding the equal sign relationally is important for success in arithmetic and algebra. Yet, many elementary school children continue to perceive it operationally, as a symbol that indicates computation. Mathematics educators have long suggested the use of the balance scale as an appropriate context for developing the relational meaning of the equal sign. However, studies that directly evaluate the pedagogical effectiveness of the balance scale remain limited. This experimental study, with 148 second- and third-grade students, examined the effects of using a virtual balance scale in promoting a relational meaning of the equal sign. Findings indicate that although the virtual manipulative may be helpful in supporting a relational meaning of the equal sign, there are other features associated with the manipulative that have advantages over the use of the dynamic virtual balance scale in promoting students' learning of mathematical equivalence. Results suggest that seemingly intuitive manipulatives that are perceptually rich may not promote optimal learning.

ACKNOWLEDGEMENTS

Many individuals and groups have supported me and contributed in their unique ways towards this dissertation project and I would like to extend my sincere thanks to all of them.

My deepest gratitude goes to my guide, Dr. Michelle Perry, without whom this dissertation would not have been possible. Michelle encouraged me to follow my hunch and heart and pushed me to stretch my limits while providing direction throughout the dissertation process. I appreciate her diligence and thoughtful reflections on my ideas and writing and her complete attention to my development as a researcher and scholar. Her dedication to the field of educational research, teaching and selfless student advising is truly inspirational. I am thankful for her time, feedback and mentorship and I am proud to be her mentee and friend.

I also want to thank Drs. Sarah Lubienski, Robb Lindgren and Daniel Hoffman for being excellent committee members. Over the past years, Sarah always found time to provide valuable advice whenever I needed it, sharing her insights on teaching and passion for the field of mathematics education. I am incredibly grateful to Robb; I took his seminar simply to explore digital learning environments and it paved way for so much more. He has inspired me to think outside of the box and offered wonderful opportunities for growth in his lab, including support to kick-start the functionality for the first prototype of my design-based project. Dan's encouragement, perspectives and thoughtful comments have been invaluable towards my work. I was fortunate to be surrounded by such knowledgeable and approachable committee members with diverse research backgrounds. I am appreciative of their support and sage advice.

I want to thank the College of Education at University of Illinois at Urbana-Champaign for supporting my doctoral studies and for funding my dissertation work through the Hardie Dissertation Award Program.

Additionally, I am thankful to the teachers who let me disrupt their class time to conduct my study, and Melissa Schmidt for her support with the data collection. And I am, of course, grateful to the many students who participated in this study and worked so diligently (and even bravely) on “strange” math problems.

I would like to acknowledge and thank my colleagues and friends in the Departments of Curriculum & Instruction and Educational Psychology for their enthusiasm and/or feedback during different phases of this study, especially: Shuai Wang, Crystal Carlson, Shereen Beilstein, Genevieve Henricks-Lepp, Jeanne Brunner, Martha Makowski, Juan Gerardo, Shana Riddick, Josef Woldense and Rob Wallon. Thanks to my very dear friends Sonya Irving Wilson and Beth Vrendenberg for always being there for me; and Stephanie Saclarides for helping me avoid turning into a social hermit, especially during the intense phases of my dissertation.

A special thanks to my husband, Gul, for his significant contribution and technical inputs in making the digital environment functional, and for making data extraction fun. I could not have completed this work without his support and patience.

Finally, I thank my parents, Balvinder and Surjit Singh Bajwa, who instilled in me values of compassion and grit as well as a love of learning from a young age and encouraged me to pursue my interests, even if that takes me thousands of miles away; and my brother, Navdeep, who is my champion, supporting me in every way possible and has believed in me for as long as I can remember. I love you dearly. You all give me the strength to move forward, even in the toughest of the times. Thank you for your unwavering support and unconditional love.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: REVIEW OF LITERATURE 5

CHAPTER 3: METHOD 29

CHAPTER 4: DATA ANALYSIS AND RESULTS 45

CHAPTER 5: DISCUSSION AND CONCLUSION 63

REFERENCES 75

APPENDIX A 84

APPENDIX B: SCREENING QUESTIONS 85

APPENDIX C: IDENTIFYING THE TWO SIDES OF THE EQUATION 87

APPENDIX D: AFFECTIVE SURVEY 90

CHAPTER 1: INTRODUCTION

The equal sign (=) is one of the central symbols in mathematics and is pervasive in topics such as arithmetic, geometry, algebra, and trigonometry. Understanding what this symbol means in these contexts is essential for success in formal mathematics. To be competent in arithmetic, for instance, children must understand that the symbol indicates a relationship of equality, symmetry, or sameness on both sides. This meaning of the symbol is referred to as the relational meaning (or relational view) of the equal sign. Unfortunately, many U.S. students across age groups continue to perceive the equal sign as an operator that indicates computation of numbers (i.e. the operational meaning) instead of as a symbol that indicates equivalence between the quantities or expressions on its two sides (McNeil, Chesney, Matthews, Fyfe, Petersen, Dunwiddie & Wheeler, 2012). Although many studies have looked at ways to help young children develop a more adequate understanding of the equal sign, it is not easy for most elementary school children to perceive the sign relationally as an indicator of equivalence.

In mathematics, educators have often utilized physical manipulatives to provide external representations of abstract ideas, which they then connect to the conventional symbolic form. A simple example of this would be when teachers use two counters to demonstrate the idea of "twoness" and connect this idea to the numeral "2" by writing it on the board. This idea is neither present in the two counters nor dependent on the objects, so it must be abstracted from the referents: the counters used to depict twoness. In other words, the meaning must be bestowed on the objects (by the instructor) and abstracted (by the learner). Even when it is abstracted, the association of "twoness" with the symbol "2" requires explicit instruction so that seeing the numeral "2" generates instantly the idea of twoness. However, abstracting meaning from physical representations or manipulatives is easier said than done. In fact, physical manipulatives

have sometimes been criticized for their potential to distract the learner from the intended meaning-making because the learner may focus on the manipulative's perceptual features (e.g., size, color or novelty of the manipulative itself) rather than the idea it holds. But recent technological advancements have reignited an interest in the potential of manipulatives for students' learning of mathematical concepts, especially manipulatives displayed on digital platforms, which are also referred to as virtual manipulatives. This is because digital platforms afford to mitigate some of the external noise associated with the handling of physical manipulatives. This interest has driven the development of several educational resources, such as the National Library of Virtual Manipulatives (NLVM, 2014) and the National Council of Teachers of Mathematics' Illuminations (nctm.illuminations. org), both of which offer access to a variety of virtual manipulatives. But while virtual manipulatives offer some advantages over their physical counterparts (e.g., accessibility), meaning still needs to be bestowed upon them through carefully crafted activities that interest students to engage in manipulation. In other words, manipulatives, whether physical or virtual, by their very nature do not speak for themselves; therefore, their pedagogical effectiveness relies in part on the context and instructional plan in which they are introduced. In fact, mathematics education policy has called for a more meaningful integration of technology into the teaching and learning of mathematics in formal situations (National Governors Association for Best Practices [NGABP], 2012; National Teachers Council for Mathematics [NCTM] 2000).

Because technology allows for the creation and control of multimedia representations, for instance on a computer, one of the ways in which virtual manipulatives can be introduced in mathematics is by embedding them within learner-friendly multimedia contexts that seek input from the user, offer immediate feedback, and guide the user to move through the manipulation

by providing relevant visual and aural instructional messages. Assuming that technology offers the potential to create new learning environments that may mitigate some of the limitations common to both physical and virtual manipulatives (e.g., variability in instructional activities involving manipulatives), it is necessary to create and test such environments to examine how students' learning may be optimized when certain perceptually rich manipulatives are used. It is also necessary to simultaneously shed light on the pedagogical effectiveness of the manipulative. In this study, I examined the pedagogical effectiveness of a virtual pan balance embedded within a multimedia instructional environment by varying the level of support offered with it and assessing how different amounts of support impact learning about the relational meaning denoted by the equal sign. I also examined how children in second- and third grade affectively respond to instruction delivered in a game-like instructional environment.

Although many acknowledge that manipulatives in general (Seo & Ginsburg, 2003; Sherman & Bisanz, 2009) and the pan balance in particular may aid learning about the equal sign (Hiebert & Carpenter, 1992; Fyfe & McNeil, 2009; Mann, 2004), it is less clear exactly which factors impact learning when a manipulative is involved. That is, although studies that have utilized a balance model imply that a pan balance scale is useful in developing a relational meaning of the equal sign, it is not known what aspect of the pan balance—the dynamic balancing of the scale, the analogical reference to balance, the crux of seeing both sides being made the same, or a combination of these—is crucial for learning about particular concepts, including the meaning of the equal sign. To address this question, I manipulated the presence of the virtual pan balance and the reference to the verbal balance analogy, keeping the visual aspects as well as the relevant audio parallel across the experimental conditions.

The central research questions this dissertation addresses are: Does a virtual pan balance scale in a multimedia instructional environment foster the development of an appropriate relational meaning for the equal sign in young learners with minimal adult direction after a brief intervention? Do different features associated with the balance scale impact learning differently when presented in the absence of the scale by offering varying amounts of support? What were students' affective responses to receiving instruction through a game-like instructional application? This study provides evidence that brief amounts of instruction can successfully cause shifts in students' understanding of the equal sign; that certain features associated with the balance scale might be more disruptive while others might be more effective in precipitating improvements in students' understandings; that more support might not necessarily be more beneficial; and that students' affect and learning are important to consider within virtual learning environments.

CHAPTER 2: REVIEW OF LITERATURE

Background

The equal sign is ubiquitous in mathematics and can be found everywhere from arithmetic and algebra to complex theorems used to devise a satellite's orbit in outer space. Having a clear understanding of this symbol is imperative for being competent in arithmetic and for successful reasoning in algebra (e.g., algebraic equations with unknowns), yet years of research indicate that many young children misunderstand this symbol.

Statement of the Problem

For many educated adults, the meaning of this pervasive symbol (=) seems simple and obvious. However, this is not the case for many elementary school children. Research indicates that many children perceive the symbol as having an *operational* meaning (i.e., "to compute" or "to find the total") instead of a *relational* one (i.e., indicating a relationship of equality between the two sides of the symbol) (Baroody & Ginsburg, 1983; Behr, Erlwanger & Nichols, 1980; Carpenter, Franke, & Levi, 2003; McNeil & Alibali, 2005a; Rittle-Johnson & Alibali, 1999). For instance, when asked to fill in the blank in equivalence problems, e.g., the form a + b = _ + c (e.g., 8 + 3 = _ + 5), children with an operational view of the equal sign tend to write 11 or 16 in the blank (e.g., Perry, Church, & Goldin-Meadow, 1988). In other words, they add all the numbers on the left side of the equal sign or add all numbers in the problem. In the past several decades, many intervention studies have been successful in helping children develop a broader view of the equal sign (Hiebert, & Carpenter, 1992; Matthews, & Rittle-Johnson, 2009; McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011; McNeil et al., 2012). Unfortunately, many U.S. elementary school children continue to have inadequate knowledge about the sign and perceive it as an operational symbol (McNeil et al., 2012). For instance, McNeil et al. found that

7- to 9-year-old children's ability to define the equal sign was poor, with only 7 of 104 children defining the equal sign relationally.

It must be noted that perceiving the equal sign as an operational symbol does not cause any difficulty while solving problems in the standard format, e.g., 3 + 4 = or 3 + 4 + 2 = _. Therefore, this problem may go undetected as students solve for one unknown by itself that appears immediately after the equal sign. However, this understanding negatively affects a child's ability to solve problems in non-standard formats involving operations on both sides (e.g., 2 + 5 = _ + 4). The operational view is especially noticeable when students are presented with non-standard problems having repeated addends with one unknown (e.g., 3 + 2 + 1 = _ + 1) and similar problems having non-repeated addends with one unknown (e.g., 3 + 2 + 1 = _ + 4). Such non-standard problems are categorized as equivalence problems and considered reliable measures for gauging young students' knowledge about the equal sign (Stephens, Knuth, Blanton, Isler, Gardiner, & Marum, 2013; Matthews, Rittle-Johnson, McEldoon, & Taylor, 2012). Moreover, children with an operational view of the symbol may only look for the operator symbol (e.g., +, -) for clues about what to do. This hinders their problem-solving ability, as they tend to perform the operation on given numbers without evaluating whether or not the answer makes sense. This problem is compounded as children move to higher grades in elementary school (McNeil, 2005; 2008). Furthermore, it acts as one of the stumbling blocks in middle school when students move to complex mathematical concepts, such as algebraic equations (McNeil & Alibali, 2005b). In fact, some evidence suggests that understanding the equal sign relationally earlier contributes to success in algebra in middle school (Matthews, Rittle-Johnson, McEldoon & Taylor, 2012). Thus, interventions that help develop an adequate understanding of the equal sign are needed.

In the remainder of this chapter I discuss the relevant literature on the concept of the equal sign, manipulatives, and multimedia learning. I will begin with a review of relevant literature on children's inadequate understanding of the equal sign. I will then highlight selective intervention studies that were successful in helping children perceive the symbol relationally. Next, I will introduce one such manipulative, the pan-balance scale, that appears in national standards (Common Core State Standards for Mathematic, 2012; NCTM, 2000), identifying its limitations as a physical manipulative and highlighting what is known and what is still missing about its pedagogical effectiveness. Finally, I discuss the potential for introducing a virtual pan balance within a multimedia-instructional environment to foster learning about the relational meaning of the equal sign.

Review of Relevant Literature on the Equal Sign

Students' Perceptions of the Equal Sign

One of the earliest studies to highlight the fact that the equal sign is greatly misunderstood across the elementary school grades was conducted by Behr, Erlwanger, and Nichols (1976). Using one-on-one interviews with first- to sixth-grade students, the researchers found that most U.S. elementary students perceived the equal sign indicating that they should "do something." For instance, in the standard format of a number sentence such as 3 + 5 =_, the equal sign was read as "a stimulus calling for an answer to be placed in the box" (p. 2). In a nonstandard sentence form, _ = 5 + 2, most of the students changed the sentence to 5 + 2 = _, mentioning that it was written backwards. In number sentences involving operations on both sides (e.g., 8 + 4 = _+ 5), students calculated the sum on left side (12) or added all the numbers present in the equation (resulting in 17). In similar problems (e.g., 3 + 2 = _ + 2), some placed an additional equal sign and changed the problem to 3 + 2 = 5 + 2 = 7. This indicated that they

completely overlooked the placement of the equal sign and focused instead on the operation presented in the number sentences. Interestingly, some students accepted the statement of the form 3 + 2 = 2 + 3 as true because the numbers were same on both sides, but they rendered it false when addends were different, e.g., in the form 4 + 6 = 3 + 7. Behr et al. noted that performance on the former statement resulted from using the commutative property of addition rather than the concept of equivalence per se. They also found that regardless of their grade, when presented with a number sentence with an equal sign, students suggested that they needed to "do something," which is an *operational* view of the equal sign instead of a *relational* one. Behr et al. discussed how exposing children to different kinds of formats in and of itself will not change this misconception, as it is a result of a "deep-seated mindset" to written number sentences (p. 10).

Many studies following Behr et al.'s study explored elementary students' views of the equal sign, and these findings supported the notion that students tend to have an operational view (Baroody & Ginsburg, 1983; Behr, Erlwanger & Nichols, 1980; Carpenter, Franke, & Levi, 2003; Freiman & Lee 2004; McNeil & Alibali, 2005a; Rittle-Johnson & Alibali, 1999). For instance, Freiman and Lee (2004) evaluated kindergarten, third- and sixth-grade students' understanding of the equal sign or lack thereof by giving equations in which blanks (i.e., unknowns) were placed either in the first, second or third position in the statements of the form $a = b + c$. The researchers also included questions in the form $a + b = c + d$, with the unknown being in the second, third, or fourth position. The third- and sixth-graders in this study performed much better than the kindergartners, which Freiman and Lee attributed to an enrichment program that the older students were participating in, which exposed them to work with different formats for a longer period of time as compared to the kindergartners. Thus, their findings highlight the

possibility that young learners may have performed better if they, too, had received enough practice for longer periods of time.

Can Elementary School Students Learn the Appropriate Meaning of the Equal Sign?

Some earlier work suggested that understanding the equal sign can come late in development, but is responsive to instruction. For example, Perry (1991) investigated fourth- and fifth-grade students' understanding of equivalence and found that almost none of these children had a relational understanding. She then provided these students with instruction involving a *principle, procedure, principle-plus-procedure,* or *procedure-plus-principle* on equations in the form 2 + 3 + 4 = _ + 4. Although instruction led to improved performance on similar problems for about half of the students across all conditions, relatively few succeeded on the transfer problems (e.g., 2 x 3 x 4 = __ x 4). However students in the principle condition performed significantly better on the transfer task than those in the other three conditions. Perry's results indicate that when students' attention is brought to the principle governing the relationship, this likely provides them a better conceptual grasp of the situation, which they can then transfer to new situations.

Rittle-Johnson and Alibali (1999) worked with fourth- and fifth-grade students, providing explicit instruction on equivalence using equations similar to those used by Perry (1991). Students were assigned randomly to one of three conditions: *conceptual, procedural,* or *control*. The pretest included conceptual and procedural tasks measuring students' understanding of equivalence. All instruction was done in a single session, followed by the posttest on the following day. The conceptual group was given equivalence problems and was asked to make the numbers on each side of the "equal sign" equal to solve the equation (similar to principle condition in Perry, 1991). The procedural group was told how to solve the equation to find the

unknown, and the students in the control group received no instruction. The posttest results indicated that the students in the conceptual and procedural conditions outperformed the control group. These findings highlight that explicit attention is helpful and leads to improvement in children's performance on equivalence problems.

Given that understanding the equivalence relationship denoted by the equal sign is receptive to instruction in older elementary school children, one may wonder whether younger children might also be able to perceive the equal sign as a relational symbol if given appropriate instruction and whether their understanding is development- or age-dependent and relatively impervious to instruction. To examine this issue, it is informative to examine children across cultures, who presumably receive different instruction, but face similar developmental constraints. Here, the evidence points to instruction as the more important factor. A cross-cultural comparison (Li, Ding, Capraro, & Capraro, 2008) indicated that sixth-grade children in higher-achieving countries performed better on equivalence tasks and generated explanations that were not strictly operational in nature, which indicates that young children *can* learn to see the equal sign relationally. However, given the age of the students, we can still ask whether instruction might be taken up successfully in younger students.

Using a similar form of equations as that used by Rittle-Johnson and Alibali (2004), McNeil and Alibali (2005) randomly assigned 7- through 11-year-old students none of whom defined the equal sign in a relational sense, to one of the following conditions: *equal-sign intervention*, *problem-structure intervention*, or *control*. The pretest involved asking students to solve an equation with operators on both sides, called "open equations" (e.g., 3 + 4 + 5 = _ + 5). The intervention presented all students with a "complete equation" (e.g., 3 + 4 + 5 = 7 + 5). *Problem structure* students received a lesson asking them to focus on the location of the equal

sign, whereas the *equal-sign* group received instruction focusing on the meaning of the equal sign (i.e., the fact that both sides of the equation need to be equal). Control group students were asked to think about the presented equation for one minute with no instruction or feedback. Results indicated that students who received the *problem-structure intervention* were able to solve the open equations, which were similar to the problems on pretest. Students in the *equal-sign* condition developed a better understanding of the equal sign as a relational symbol. Students' dependence on procedural strategies on the pre- and posttests suggested that traditional instruction interfered with their performance post-intervention. McNeil and Alibali also suggested that students were more likely to learn from the treatments if they had not already mastered procedures.

The above studies imply that young children are capable of perceiving the equal sign as a relational symbol and that they should be introduced to this meaning sooner rather than later. Such findings naturally beg the question: if elementary students can learn to see the symbol relationally, what could possibly be causing them to form an operational meaning in the first place? This is crucial if we are to design instruction that interrupts or undoes students' misunderstandings.

Possible Reasons for U.S. Elementary School Students' Inadequate Understanding of the Equal Sign

As mentioned earlier, many studies identified that the difficulty of understanding the equal sign can be attributed to the way the symbol is first introduced in elementary school, by using standard-format addition problems (Baroody & Ginsburg, 1983; Capraro, Ding, Matteson, Capraro & Li, 2007). Practicing arithmetic almost exclusively in standard formats, where the unknown always appear alone and on the right side, leads to an implicit understanding of the

equal sign as meaning "to compute" or the "answer comes next" (Behr, Erlwanger & Nichols, 1980; Capraro et al., 2007; Carpenter, Franke & Levi, 2003; Kieran, 1981; McNeil & Alibali, 2004; McNeil et al. 2006; McNeil et al., 2011).

Another reason for perceiving the equals sign as an operational symbol can be attributed to the fact that little to no explicit attention is given to equal sign instruction in elementary school mathematics. In a comprehensive study evaluating U.S. curricula designed for grades K-5, Powell (2012) found that seven out of the eight curricula did not include non-standard equations, which can help foster a relational meaning of the equal sign. She also found that the equal sign was rarely mentioned and not given explicit attention. Although there were some references to the concept of equality denoted by the equal sign in the kindergarten and first grades, references to equality were mostly missing from second grade onwards, even though the symbol appeared frequently. Additionally, only half of the curricula had pictures or icons or suggested using a manipulative (e.g., balance scale) to encourage the development of an understanding that both sides of an equation are the same, and this use was not consistent.

In a similar study comparing teachers' mathematics education books in China and the United States, Li, Ding, Capraro, and Capraro (2008) identified the lack of explicit attention to the equal sign as a possible reason for misconceptions among U.S. students: "Out of the six U.S. textbooks, only two directly addressed the equal sign, and none includ[ed] lesson examples or activities to help understand how the equal sign should be taught" (p. 209). In contrast, the Chinese elementary curricular materials, including teachers' guidebooks, included materials that varied in *context*, which were designed to develop an understanding the equal sign. The guidebooks emphasized the relational meaning and advocated for using one-to-one correspondence to teach this concept. In the case of arithmetic, tasks involving equations were

given in a variety of contexts to foster an understanding of the equal sign as "the same as" or "balance." Additionally, one of the units explicitly stated "that the goal of the unit is to learn and understand the "=", "<" and ">" (p. 204), thus indicating special and explicit attention to this symbol in comparison to other relational symbols. Li et al. suggested that U.S. students' narrow view of the equal sign and consequent difficulty in equivalence problems might be attributed to the stark contrast in specific attention given to understanding the equal sign in U.S. versus Chinese curricular materials. In fact, Hattikudur and Alibali (2010) found that teaching the equal sign along with other signs of inequality to third- and fourth-grade students promoted their conceptual understanding of the equal sign. Thus it appears that young U.S. students predominantly have access to standard-format arithmetic problems, where any misconception of the equal sign cannot be detected or remedied. In this case, many U.S. students become entrenched in their (mis) understanding. Given that this problem has been identified, it is important to understand what sorts of interventions might foster a relational understanding of the equal sign and how early can this be delivered to young students?

Interventions to Help Elementary School Students Develop an Appropriate Understanding of the Equal Sign

Baroody and Ginsburg (1983) conducted a year-long study involving an intervention with first- and second-grade students to develop a relational meaning of the equal sign. The teachers always referred to the equal sign as meaning "the same number" whenever it appeared, for the span of an entire year. The researchers showed students equations without an unknown and asked them whether or not they were acceptable. Some of these equations were similar to ones they had seen in the year-long intervention (e.g., $5 + 6 = 12 - 1$); others were unfamiliar equations in the form $x = x$ (e.g., $5 = 5$). At the end of the year, first-graders performed better

than second-graders. The authors suggested that the better performance of the younger students may be attributed to the fact that they had only practiced the new curriculum, which involved exposure to different formats and attention to the equal sign meaning, whereas the second-grade students had had a full year of traditional curriculum instruction, and during that time they did not have any focus on the relational meaning of equal sign. This full year of practice with arithmetic without attention to the relational meaning of the equal sign may have entrenched them in their operational views and interfered with the second-grade students' learning during the intervention. In any case, this research suggested that students *can* learn a more appropriate understanding of the equal sign (i.e., the relational view) even in first grade as a result of focused and prolonged instruction. The limited transfer in this study indicates that although using the verbal equivalent of the meaning of equal sign (e.g., "the same number") has potential, it does not indicate if transfer to different formats with unknowns is feasible as a result of such instruction. To examine this issue further, the current study will evaluate whether or not transfer to novel formats (i.e., other than those used in the instruction) involving a single unknown is possible when the equal sign is associated with the verbal message "the same as."

As mentioned earlier, one of the ways that students' perception of the equal sign as a relational symbol can be promoted is by providing exposure to different equation formats in arithmetic (i.e., standard as well as non-standard formats with an unknown placed in different positions). Some studies support the idea that exposure to non-standard formats with one unknown can help children shift their understanding of the equal sign from an operational to a relational meaning (McNeil et al., 2012; McNeil et al., 2011). However, it has also been found that students do not always respond to instruction that focuses on developing a relational meaning of the equals sign. For example, in some studies (McNeil, 2005b; 2007) the researcher

found that first through fourth grade students respond to instruction on the relational meaning of the equal sign differently and students' performance to equivalence problems declines between the ages of 7- 9 years, with 9 year olds most strongly resisting an instructional change that focuses on altering their deep rooted operational patterns. In another study with third graders (Molina & Ambrose, 2006), when children learned the relational meaning of equal sign, they did not retain this understanding. The researchers found that third-grade students' immediate posttests reflected that students had learned the relational meaning; however, only 5 of 15 students could identify the statement 3 = 3 to be true, and only 10 students could identify 7 = 12 as false. These results further support the idea that causing students to completely alter their notions about the equal sign from an operator to a relational view may not be achieved easily.

In fact, narrow early experiences with the equal sign in arithmetic have been found to be long lasting and continue even at the undergraduate level (McNeil & Alibali, 2005b). For instance, McNeil and Alibali found that some high school and undergraduate students continue to have these misconceptions around the meaning of the equal sign. They found that when operational patterns were activated, even undergraduates resorted to an operational view.

Given that the operational view may persist for a very long time, it is striking and important that students can potentially benefit from instruction as early as first grade (e.g. Baroody & Ginsburg, 1983; McNeil & Alibali, 2005). It is possible that if this misconception can be addressed early on, related difficulties in arithmetic (McNeil et al. 2006; Molina & Ambrose, 2006) and advanced mathematics (e.g., algebraic equations) can be avoided. McNeil et al.'s studies validate the need to attend to the equal sign early on in schooling, as mere exposure to non-standard formats at a later stage may not help in reverting years of holding a misconception (McNeil, 2005). In addition, given the potential significance of developing a relational

understanding early, attention should also be devoted to alter this misconception effectively in students who have already been through some amount of formal schooling. The current study includes students from similar age groups as those of McNeil (2005a, 2005b) and examines if practicing equivalence problems together with some amount of external support impacts students' learning and, if so, whether more or less external support is beneficial in breaking the existing operator view.

In the next section, I will highlight selective studies that have explored the use of manipulatives in supporting a relational view of the equal sign.

Using Manipulatives to Promote a Relational View of the Equal Sign

Aside from exposing children to different formats, a few studies have looked at how to introduce the concept within a context by utilizing manipulatives. Manipulative use has been promoted in the curricular materials of high achieving countries, as found in international comparisons of texts and curricular materials (Li, et al., 2008). As an example, Seo and Ginsburg (2003) noted that second-graders were inconsistent in exhibiting an operator versus a relational view. In particular, in concrete scenarios without any operations (i.e., 1 dollar = 100 pennies or one red rod = 2 white rods), most students responded with a relational description of the equal sign. But when presented with number sentences without a concrete context, both in standard (i.e., $a + b = c$; $a - b = c$) and in non-standard form (e.g., $c = a + b$; $a = a$), most students responded with an operational description of the equal sign. Seo and Ginsburg suggested that the lack of context promotes an operational view, but real-life and concrete, non-symbolic contexts promote a relational view.

Sherman and Bisanz (2009) examined the same issue by evaluating second-graders' (n = 48) accuracy on non-symbolic and symbolic problems. They designed a non-symbolic method to

determine if second-graders' misunderstanding of the concept of equivalence was a result of the presence of symbols or due to children's inability to comprehend equivalence per se. They tested how two groups of participants performed on equivalence problems with two addends on both sides of an equal sign, using non-symbolic and symbolic representations of the problem with students. The non-symbolic problems presented to the students relied on manipulatives to represent the quantities in an equation (i.e., blocks to represent addends and plastic bins to hold each addend). The symbolic problems were presented on a worksheet, in the form $a + b = c + _$ (where a, b, c, d were all positive integers). Participants were asked to explain why their answer was correct, and feedback was provided. Participants in the non-symbolic condition gave more justifications—about three times as often as students in the symbolic condition. Also, Sherman and Bisanz noted that children in the non-symbolic condition were more likely to give relational responses than their counterparts in the symbolic condition. They further tested if presenting non-symbolic problems first would impact students' accuracy on subsequent symbolic problems, mapping their understanding of equivalence on the equal sign. The results indicated that children who solved non-symbolic problems first had much greater accuracy on symbolic problems than children who solved symbolic problems first. Additionally, experience with non-symbolic situations involving manipulatives benefited their subsequent performance on symbolic problems. These findings highlight the potential benefits of using manipulatives in addressing students' misconceptions of the equal sign as well as manipulatives as a means to bootstrap this learning to solving equivalence problems using conventional notation. This work further grounded the notion that young children can understand equivalence based on the context in which the symbol is introduced and understanding equivalence is not maturational but instead can be influenced by situating learning in concrete contexts. The findings also indicate, but do

not decisively answer, whether associating the symbolic conventional symbol with meaningful concrete experiences might prove useful to support development of a relational understanding. What remains unanswered is whether situating the association of the symbol within a context also helps in breaking previously formed operational patterns? This issue was one that was addressed in the current study by working with second- third-graders who demonstrated operational understandings of the equal sign.

Using the Balance Scale for Developing a Relational Meaning of the Equal Sign

The above cited work shed light on the potential of a meaningful context that provides concrete experience to children to understand equivalence, and in a similar vein, researchers have indicated that effectiveness of manipulatives lie in part on how well suited the concrete model is with the idea the model is being used to elicit. Hiebert and Carpenter (1992) argued that the closer the match between the characteristics of the manipulative and those of the mathematical relationships, the more conceptual support students have available to construct the intended connections. For this reason, the balance scale appears to provide an excellent *context* and manipulative, as the act of balancing the two sides is dependent on them having the same amount (in weight), which provides a conceptual correspondence to the relational meaning of the equal sign. In fact, Hiebert and Carpenter (1992) suggested that associating the fulcrum of the balance with the equal sign could act as a model for developing a relational understanding of the equal sign. Perhaps the fact that a balance scale, by its very nature, offers an external model of equivalence has been the reason why this manipulative has been emphasized in national policy documents (NCTM, 2000; CCSSM, 2012), some pre-service teacher guides (e.g. Baroody & Coslick, 1998; Van de Walle, 2012), and research-based elementary school textbooks (e.g., *Everyday Mathematics*). For instance, the Common Core State Standards for Mathematics

(CCSSM, 2012) have reiterated the importance of algebraic thinking in grades K-5 to support students' future learning in algebra in later grades and suggest "using a pan balance." However, they do not explain how to use this manipulative. Surprisingly, in spite of years of claiming its potential in developing algebraic and pre-algebraic understanding (Baroody & Coslick, 1998; Hiebert & Carpenter, 1992), as well as the fact that the balance scale is recommended in 50% of the teachers' manuals accompanying popular U.S. elementary school mathematics textbooks [4 out of the 8 analyzed in recent years by Powell (2012) and many teacher education books (e.g., Van de Walle, 2012; Baroody & Coslick, 1998)], there is little empirical evidence on what factors in the balance scale manipulative promote learning.

Some classroom-based work (Mann, 2004) indicated positive results when it comes to using the concept of balance in mathematics instruction. For example, Mann had her students use their bodies (e.g., stretched arms) for balancing and found improvement in students' levels of comfort with arithmetic problems. Mann (2004) indicated that when second-grade students were introduced to the equal sign in the context of a seesaw, they could come up with algebraic rules. Emphasizing the balance properties inherent in the seesaw analogy, the teacher encouraged her students to use their hands as the two ends of the seesaw and then asked them to guess what would happen if she would put different combinations of heavy and light objects on each side. Seeing the effects of putting and taking away different objects on the two ends of the seesaw made with objects and their bodies, students were encouraged to come up with "seesaw rules" that were later narrowed down by the class. At a later time during the year, students were given challenging questions (e.g., $a + b = _ + c$) and encouraged to use their knowledge of the seesaw to check their solutions. This classroom-based work indicates the potential of using the balance analogy in building a conceptual understanding of the relational meaning and the rules associated

with equality in elementary school children. However, because it was part of a whole year of instruction and was possibly coupled with other instructional activities, the effectiveness of the balance analogy or the use of the seesaw as balance representation was not tested specifically. Hence, the effectiveness of the balance model cannot be ascertained and remains an open question.

Findings from recent work have also demonstrated the potential in using manipulatives, particularly for learning and transfer with a "concreteness fading" approach (McNeil & Fyfe, 2012; Fyfe, McNeil, Son, & Goldstone, 2014). Specifically, Fyfe and McNeil (2009, unpublished work) evaluated a three-step-progression approach for teaching equivalence to young children. Using Bruner's three-step model of instruction (concrete-iconic-abstract representation), the authors tested whether providing experience with concrete manipulatives (e.g., enactive experience using a physical pan balance), followed by a gradual removal of the concreteness, led to learning and transfer. The participants in the concrete-to-iconic-to-abstract condition (*concreteness fading*) condition performed better on transfer items than did any of the other groups. However, the authors suggest that further research needs to be done to identify if certain steps in concreteness fading act as a moderator and are essential in learning about a concept, while others can be removed.

In summary, recent work provides some indication that a balance scale might be an effective manipulative for supporting learning and transfer of mathematical equivalence. Less clear is what aspects of this manipulative—the dynamic visual from unbalanced to balanced scale, or providing a verbal description of balancing, or focusing on the two sides being the same in quantity, or a combination—are crucial for learning. Moreover, we do not know whether a brief intervention that utilizes some external support using manipulatives can lead to a shift in

students' perceptions from the operational to the relational view. In the current study, I addressed these gaps by examining if brief amounts of instruction given with minimal adult direction within a game-like instructional environment can precipitate changes in students' thinking and, if so, whether the different amount of support offered in the instruction impacted learning differently.

Affordances and Constraints of Physical Manipulatives

In general, the utility of a physical manipulative lies in its ability to offer a concrete form of external representation to scaffold understanding of abstract concepts (Ainsworth, 2006). Additionally, physical manipulatives may provide a practical context useful for activating either physical or mental action or for enhancing memory and recall (Glenberg, Gutierrez, Levin, Japuntich & Kaschak, 2004). Various national documents for mathematics instruction emphasize the potential of using physical manipulatives to foster mathematical learning and reasoning throughout the grades (NCTM, 2000; CCSSM, 2012).

It is widely acknowledged that the use of physical manipulatives does not come with guarantees that learning will occur (Ball, 1996; Baroody, 2006; Clements & Sarama, 2004), as meaning has to be bestowed upon the objects, and this is not always easily achieved. On the contrary, working with physical manipulatives may actually cause the learner to get distracted by the extraneous perceptual and motor details of the manipulative (Kaminski & Sloutsky 2013), which may draw the learner's attention to the manipulative itself rather than to its referents (Uttal, Scudder, & DeLoache, 1997). This may tax the learner's working memory, causing cognitive overload (Mayer & Moreno, 2003) and thus leading the learner not to engage in the targeted meaning-making.

Besides these difficulties in relying on physical manipulatives to support the development of new knowledge, we can also hypothesize that some of these manipulatives might be difficult to procure and, even when procured, may be cumbersome or messy to use. For instance, using a physical pan balance requires careful handling of the apparatus, which involves two pans and weights, and selecting weights that do not roll within the pan is necessary to avoid shifts in maintaining balance, which may cause measurement errors. In instances when only one balance scale is available and is used for demonstration, the perspective of the viewer may also become important, and the teacher need to make sure to refer to the right side and the left side so that all students follow accurately. Additionally, building conceptual knowledge, using most manipulatives in general and using the pan-balance in particular, requires a lot of back and forth across the physical manipulative (the scale), the abstract idea (balance), and the symbolic representation (the equation). Kaput (1989) has discussed this as a common problem in using physical manipulatives, emphasizing that the connection between the actions on the manipulatives and those on the symbols are not easily perceived.

Using virtual counterparts of physical manipulatives can minimize some of the above-mentioned limitations of the physical manipulatives because many such messy situations can be controlled. Next, I highlight how virtual manipulatives, in general, and the virtual pan balance, in particular, helps mitigate some of this messiness, making a case for the use of a virtual rather than a physical manipulative for the purposes of this study.

The Pan Balance as a *Virtual* Manipulative

In the previous section, I noted the benefits and pitfalls of using an actual pan balance to support students' understanding of the relational conception of equivalence. In the past decade, however, great strides have been made in utilizing technology to promote learning in both formal

and informal settings (Clark & Mayer, 2011; Mayer, 2009). In mathematics, a result of this has been an increasing interest in creating and including *virtual manipulatives*, instructional applets, and computer-based tools as part of instructional materials. In fact, the National Library of Virtual Manipulatives (NLVM, 2014) and the National Council of Teachers of Mathematics (NCTM illuminations) have reservoirs of virtual manipulatives available on their websites.

Virtual manipulatives are "computer-based renditions of common mathematics manipulatives and tools" (Dorward, 2002, p. 39). Although there are a variety of ways in which virtual manipulatives are defined, minimally, virtual manipulatives are visually similar to their physical counterparts (Moyer, Bolyard & Spikell, 2002; NLVM, 2014) and are more than static images. Also, virtual manipulatives at the bare minimum have added affordances as compared to their physical counterparts. One of the affordances is a cleaner representation and smoother functioning of the manipulative. Additionally, the possibility of providing immediate and consistent feedback can be more easily achieved in a digital platform than in an actual classroom because the platform responds to each individual student's error.

In the case of a pan balance, some of the limitations or "noise" associated with the concrete pan balance can be controlled in the virtual medium. For instance, the messiness associated with maneuvering a physical pan balance can be mitigated easily with a virtual pan balance and measurement errors resulting from the sensitivity of different scales can be controlled. Additionally, the presentation of the balance scale and the corresponding symbolic notation can occur simultaneously or closely in time and visual space, thereby linking the demonstration of equivalence with the equal sign symbol. This diminishes the distractions of the back-and-forth nature of the physical manipulative. In general, many extraneous factors can be controlled, and the presentation of information can be manipulated more efficiently in a virtual

medium using a virtual manipulative than with the actual manipulative. Furthermore, the symbolic form can be closely linked and presented with the actual balanced state of the balance scale on the screen simultaneously and with appropriate aural feedback.

In sum, a virtual environment provides the affordances of presenting instructional steps precisely thereby allowing students to focus at particular information at a given time. Working on a computer screen also allows the *visual* and *symbolic* messages to be presented without considerable time lag, something that is not easily achieved when using a physical pan balance. This not only reduces the cognitive load but also makes it possible to highlight the key relationship or the idea underlying the model of a pan balance.

In fact, Clements (2002) suggested that well-designed computer programs might be better than physical manipulatives because a virtual manipulative allows for "reflexive thought" because it links the action explicitly to the symbolic format. As such, it is necessary to examine how technological benefits can contribute to enhance learning experiences by mitigating limitations of physical manipulatives and, furthermore, how best to provide instructional support for students' robust learning with virtual manipulatives.

As mentioned earlier, prior interventions on developing an understanding of the equal sign have primarily utilized traditional forms of instruction; therefore, augmenting instruction by using technology has been little explored. In the next section, I will highlight the affordances of technology for delivering instruction for developing a relational meaning of the equal sign, specifically to address the research questions investigated in this study.

Promise of Technology: Structured Discovery Using an Instructional Multimedia Environment

Although the virtual counterpart of a pan balance scale, or physical manipulatives in general, offers a cleaner version of the physical manipulative by controlling some of the extraneous factors that may hinder learning, the success of using any manipulative largely depends on how it is used in the classroom by teachers and students (Moyer-Packenham, 2010). In fact, research suggests that the teacher effect is responsible for the largest amount of identifiable variance; this is true regardless of students' grade level or the difficulty of the mathematics topic. One affordance of technology is that it can be done with minimal adult direction, thereby mitigating the variance associated with different instructors or even with the same instructor across multiple students. In other words, technology can standardize the presentation of material. Moreover, virtual manipulatives allow learners to move forward at their own pace.

Technology also allows the designer to choose what messages are presented in which input channels. In general, multimedia presentations are often seen as more beneficial than single media presentations (Mayer, 2001), because they both engage learners and support the development of representations. Moreover, because each information-processing channel is limited in capacity (Chandler & Sweller, 1991), an overload of information through any channel increases cognitive load and negatively affects learning. Proponents of multimedia learning argue that information that is presented in both auditory and visual forms in close proximity supports learning by reducing the cognitive load in working memory (Mayer, 2005c; Clark & Mayer, 2011). Virtual manipulatives situated within a meaningful instructional context can thus enhance learning by offering multimedia messages (Clark & Mayer, 2011; Mayer, 2009).

Learners are more likely to be engaged when the content is delivered meaningfully within a context that is relevant to the learner's age; this is especially true in the case of young children. Research indicates that young children are naturally inclined to play and there is some evidence that indicates that playing games can provide opportunities for children to develop number-sense in a meaningful and engaging way (Ramani & Siegler, 2008). There is also some evidence that young children can benefit from a wide variety of well-designed and motivating computer games (Clements & Sarama, 2002, 2004). Because the intervention in this study was designed to be instructional in nature, in order to offer a child-friendly context, the interventions were situated within a game-like context within a multimedia environment on the computer.

As mentioned earlier, virtual manipulatives have a variety of advantages over physical manipulatives and can thus aid learning as well as teaching (National Mathematics Advisory Panel [NMAP], 2008). This is especially true for linking the concepts with their symbolic notations. One can find a surge in virtual games and manipulatives, but there is a dearth of studies that systematically test which aspects of a manipulative support learning, and examine the pedagogical effectiveness of specific manipulatives (e.g., pan balance; T-balance; base-ten blocks).

Summary and Purpose of the Study

The review of studies on the equal sign indicates that students benefits from explicit instruction on the relational meaning of equal sign (e.g.,McNeil & Alibali, 2005; Rittle-Johnson & Alibali, 1999). Early intervention in elementary grades may help in overcoming misconceptions about the equal sign (Baroody & Ginsburg, 1983; Falkner et al. 1999). Additionally, for students to understand the meaning of the equal sign and to strengthen their sense of equivalence, it is important to introduce the symbol and the concept of equivalence

within meaningful contexts (Baroody & Ginsburg, 1983; Sherman & Bisanz, 2009) that can be mapped to the formal symbolic notations (Hiebert & Carpenter, 1992).

Although many studies have advanced the field in terms of instructional changes that can foster learning of the concept of equivalence, most of these interventions have focused on traditional forms of instruction, especially direct teacher instruction. Given the affordances of recent technological advances, which minimize cognitive overload and can place instruction and learning in game-like contexts, it is appropriate to ask whether a pan-balance game might support young students' learning about the relational meaning of the equal sign.

Thus, the purpose of the present study was to utilize the affordances of technology to help young students understand a relational meaning of the equal sign. I wanted to evaluate the effects of a balance analogy in a) promoting an explicit understanding of the relational meaning of the equal sign, b) improving success on equivalence tasks, and c) promoting transfer to unfamiliar problem structures in second- and third-grade students who, before any intervention, rely on the operator view of the equal sign. Additionally, I wanted to examine which aspects of the analogy and manipulative in a virtual environment are most likely to support student understanding of equivalence.

Significance of the Study and Research Questions

This study contributes to the literature and interventions focused on improving students' understanding of the equal sign and consequent performance on equivalence problems by designing a multimedia instructional intervention that addresses this concern. Empirically, this project examined the pedagogical effectiveness of a balance scale within a multimedia game-like instructional application. Theoretically, it sheds light on the amount of support inherent in the

manipulative, and whether learning is impacted differently with different amounts of support on a digital platform.

Specifically, the study investigated the following research questions:

1. Can brief amounts of instruction within a game-like instructional environment that uses some amount of external support cause shifts from an inappropriate operational understanding of the equal sign to an appropriate relational understanding?
2. Does using more support (i.e., a dynamic balance scale) or lesser amounts of support (e.g., a balance analogy without the dynamic scale) in the brief instructions within the game-like environment impact learning differently?
3. Is there a significant drop-off in learners' understanding one week after the intervention they received?
4. What are students' affective responses to the instructional game and do these differ by the amount of support?
5. Will students' judgments of learning reported during the survey map onto their actual performance on equation solving at the immediate posttest?

CHAPTER 3: METHOD

Introduction

This study was designed to evaluate what features associated with the balance scale offer optimal support within a game-like instructional environment to young learners. Additionally, the study explored how students responded affectively within a virtual game-like environment across these varied levels of support. In this case, the environment focused on helping learners develop an understanding of the relational meaning of the equal sign. Specifically, the study sought to examine whether learners understood the equal sign's relational meaning better when exposed to instruction that uses: (1) the core idea of two sides being the same in quantity, (2) the core idea coupled with a sense of those quantities being balanced with one another, or (3) the combined core idea and balance analogy overlaid with a virtually concrete balance scale (4) when compared to a control. In other words, I aimed to evaluate what levels of support associated with a balance scale impacted learners' understandings of equivalence as denoted by the equal sign.

Setting and Participants

A total of 171 second- and third-grade students enrolled in four public elementary schools in the Midwestern United States participated in the study. Of those students, 20 students were at the ceiling or near ceiling at the pretest; two were removed from analysis due to experimenter error, and one did not complete the intervention. Thus, a total of 148 participants were considered for analysis and had the mean age of 8.4 years (S.D = .73). Of these, 56% female, approximately 52% were White, 27% Black, 9% Hispanic, 1% Asian, and 11% multi-racial. 77% of participants were eligible for free and reduced lunch; 10% were second English language learners; and 10% were on an individualized educational plan.

I targeted the second- and third-grade students because studies have indicated that by second-grade children show difficulty in altering any pre-formed misconceptions regarding the equal sign, and resist a change towards forming a relational understanding (McNeil, 2005a), which is an important pre-algebraic concept (Baroody & Ginsburg, 1983; Knuth, et. al., 2006).

Research Design

This experimental study was conducted with the independent variable as one of the four type of instruction students received on equivalence problems in a virtual game-like instructional environment: (a) control, (b) core-idea, (c) verbal analogy or (d) balance scale, described under materials. Students' performances on three components of the equal sign understanding served as the dependent variables. After excluding students who were at or near ceiling, based on their performance on an equation-solving task (solving at least 5 of the 6 problems correctly) during the pretest (n=20), there were 40 students in the Balance Scale condition, 38 students in Analogy condition, 38 students in the Core-Idea condition, and 32 students in the control condition.

Instrumentation

The three components of equal sign understanding used in this study were: (1) students' performance on solving equivalence problems with one unknown, (2) students' definition of the equal sign, and (3) students' understanding of the structure of an equation by identifying the two sides of a given number sentence. These measures capture related but different components of students' understandings related to the equal sign, and have been validated in prior studies as reliable measures (Knuth et al., 2006; McNeil & Alibali, 2005; Perry, 1985, 1988; Rittle-Johnson, et al., 2011). A screening test captured students' competency in finding basic sums and differences and true-false questions on determining whether a given number sentence was true and was the same as used by Rittle-Johnson et al, 2011. I designed an affective survey that was

administered post-intervention to capture students' impressions of the game-based environment, including how often they play on computers or a laptop during a week. Details on the instruments are provided under materials.

Procedures

Data for the current study were collected in the 2014-2015 school year. Each student participated in a total of three sessions on three separate days. With the exception of the first session, which was administered to the whole class and took place in the classroom, students were seen individually outside of their classrooms in a school-designated area, during regular school-hours, either by the researcher or by a trained research assistant (collectively referred to as trainer from hereon). The intervention and all instruments, otherwise noted, were administered using Apple MacBooks. The affective survey and the third component of identifying two sides of the equation were paper-based.

First Session

Screening test: All participants with parental consent took a paper-and-pencil test, with 8 true-false problems, 4 standard addition, and 4 subtraction problems as part of the screening test. This test was administered in classrooms and lasted approximately 30 minutes. Participants who demonstrated at least 25% success on addition and subtraction problems combined were included for further participation. Students from each classroom were randomly assigned to the four conditions based on the screening scores.

Second Session

From the second session onwards, each student met individually with a trainer. The second session had four phases: 1) the pretest, 2) the intervention on the computer, 3) the

immediate posttest and 4) an affective survey that was read to the students, given in that order. This session lasted approximately 60 minutes, including a brief break.

Third Session

To gauge whether there was any significant drop off in learning, a delayed posttest, identical to the immediate posttest, was administered one week after the second session. The trainer prompted students to think of the instructional game they had played the last time the trainer saw them, and asked them to solve problems in the same way that they had done previously. The delayed test was identical to the immediate posttest.

Materials

Assessment materials included a screening test, a pretest, an immediate posttest, a delayed posttest, and an affective survey instrument. The intervention materials for this study consisted of four versions of an instructional applet.

Assessment

Screening test. The screening test was comprised of four addition problems and four subtraction problems in standard formats, with a blank after the equal sign (e.g., 4 + 5 =_). The purpose of screening students was to rule out the possibility that a failure to perform on equivalence tasks was a result of a lack of conceptual understanding of addition or subtraction. Students were asked to solve the problems as best as they could and write their answers in the blanks. Additionally, six true-false problems required students to decide if the given number sentence was true (see Appendix B). These problems entail understanding of the equal sign and are arguably introduced in the first grade (CCSSM, 2012).

Pretest, immediate posttest and delayed posttest. The pretest, immediate posttest and delayed posttest consisted of three components designed to capture students' understandings of

the equal sign. Students' response times were logged through the computer's internal timer clock. Next, I will describe the specifics of the pretest and the posttest. The delayed posttest was identical to the immediate posttest and hence is not described separately.

Solving equivalence problems with one unknown. Students were presented with six equivalence problems with one of the addends repeated, one at a time on the computer at the pretest. Once they reported the number that they thought should go in the blank, the trainer followed up by asking, " Can you tell me how you figured that out?" Students' responses were noted.

At the posttest, students were presented with six equivalence problems similar to the pretest (practice) but were also presented with three transfer items (see Table 1). These transfer items did not have a repeated addend but required the fundamental understanding that the sum on both sides of a given problem had to be made the same.

Table 1

Solving Equivalence Problems

Problem Type	Pretest	Posttest
Practice	2 + 5 = 4 + _	2 + 4 = 5 + _
	3 + 6 = _ + 8	3 + 5 = _ + 7
	7 + 2 + 5 = _ + 5	7 + 2 + 4 = _ + 4
	3 + 2 + 5 = 3 + _	5 + 4 + 3 = 5 + _
	4 + 2 + 7 = _ + 7	8 + 2 + 4 = _ + 4
	5 + 3 + 4 = 5 + _	2 + 3 + 5 = 2 + _
Transfer	Not applicable	6 + 2 + 3 = 7 + _
		8 + 3 - 2 = 8 + _
		5 + 1 + _ = 6 + 2

Explicit definition of the equal sign. To assess students' explicit understanding of the meaning of the equal sign, prior to intervention, students were presented with the equal sign symbol on the computer screen. The experimenter pointed to the screen and asked the following questions, each separated by the student's response that was entered into the computer and/or the trainer's recording sheet: (a) What is the name of this symbol? (b) What does this symbol mean? (c) Can it mean something else? Following the last question, students were presented with a paper and asked (d) If you were to show or explain to a friend or someone who doesn't know what the equal sign means, how would you explain it to them? Could you draw or write something that you would use so they could understand what the equal sign means? The first three questions are based on those used by previous researchers (Baroody & Ginsburg, 1983; Knuth et al., 2006.). I added the last question to elicit any other potential understanding of the equal sign.

At the posttest, to examine if students had any shifts in how they defined the equal sign after the brief intervention, they were asked: "After playing the game, do you think the equal sign can mean something else?"

Identifying two sides of an equation. Two questions, presented on paper, and adapted from previous research done by Rittle-Johnson and colleagues, assessed students' perceptions about breaking a given number sentence into two distinct sides separated by the equal sign. Students were shown five possibilities that separated the two sides of each of the number sentences (see Appendix C). The trainer prompted students: "Here is a number sentence and it has two sides." The trainer then pointed to each of the choices that depicted the possible separation of the two sides by a vertical line and said, "Some students thought this is how we should break the number sentence into two sides, so this (circling the side to the left of the

vertical line) is one side and this (circling the side to the right of the vertical line) is the other side. Can you take a look at all of these choices and circle the choice that you think breaks the problem correctly into two sides?"

Affective survey. An affective survey was given at posttest only and was read to the students one question at a time (see Appendix D). The purpose of the affective survey was to examine students' affects to receiving intervention on a digital platform and their judgments of learning after the intervention.

Intervention Materials

A game-like virtual environment (Marty's Numberland) was designed by the researcher to examine the effect of different levels of support associated with the balance scale on students' understandings about the equal sign. The design elements utilized the contiguity and modality principles of multimedia learning (Clark & Mayers, 2011; Moreno & Mayer, 1999). An animated character (cat) voiced the narration and appeared for story-building purposes and for providing feedback. All participants practiced six problems within a game-like environment on the computer screen and used a mouse-clickable interface to proceed in the environment. The different levels of support were represented by four different versions (applets) of the environment. Game features such as the central character were maintained across instructional conditions to rule out novelty effects or affective responses resulting from presence of a virtual character.

Instructional applets. The **"control"** instructional applet offered no visual support beyond presenting students with symbolic problems in the virtual environment. The next level of support, the **"core idea"** instructional applet, involved the task of making the two distinct sets of virtual blocks the same. The **"verbal analogy"** instructional applet, instructed students to make

the two sides, represented by two separate pans, balanced so they became the same. The "**balance scale"** instructional applet was designed to offer the most support. Students in this condition were instructed to make a virtual pan scale balanced. In other words, the levels of support occurred on a continuum from *least to most* in the following order among the instructional conditions: control, core idea, verbal analogy, and balance scale. Next I describe the progression of the environment, including specific differences among the four applets during the introduction phase and play phase.

Game introduction. Once the student put on the headset and the environment was launched, the animated character appeared and was introduced by the voice narration in the game as a smart cat that likes to play with numbers. After this, students in the verbal analogy and balance scale condition saw a static pan balance scale graphic as part of the introduction (*Figure 1*).

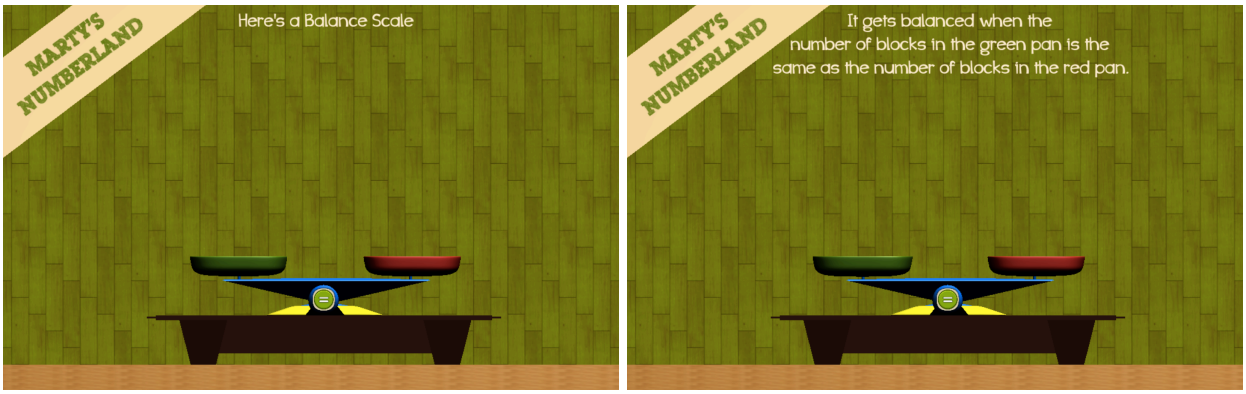

Figure 1. Introduction to the balance scale

The narration continued with, labeling of the left pan as Marty's and the right pan as the student's. The animation then dropped one block on the left side and three blocks on the right side, replacing the green equal sign on the fulcrum with a lock, accompanied with relevant narration. Following this, the text and accompanying voice-over asked the student "How will the scale move?" The student was prompted to click on one of three options that indicated what

would happen to the scale (*Figure 2*). This prompt and introduction was to initiate a balance analogy and get the student familiar with workings of a balance scale.

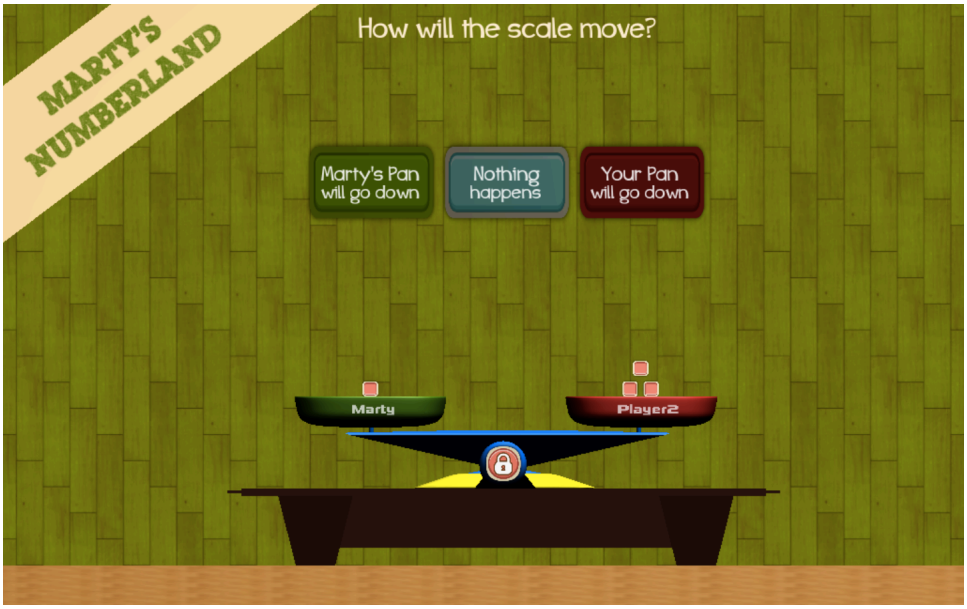

Figure 2. Initiation of the balance analogy

Because the study aimed to control for levels of support, certain features were controlled among applets. For example, in the introduction (and during the play phase) when participants from the maximum support group (balance scale) answered correctly, (1) the balance scale animated such that the heavier pan went down, (2) the lock icon displayed at the fulcrum was replaced with a not-equal symbol (≠), and (3) the narration explained that the scale is not balanced as the green pan does not have the same number of blocks as the red pan. Followed by "this side (zooming in on the left side) is not the same as this side (zooming in on the right side)". Students in the verbal analogy condition received similar narrative feedback during the introduction, but did not see the scale move. In this condition, the scale was replaced with two empty pans during the play phase. Students in the core idea group and the control group were not introduced to the balance analogy and did not see the pan balance scale graphic but received parallel relevant introductions and instructions.

Play phase. After the introduction**,** the play phase started with "Let's play with Marty" and the voiceover introduced the mission of the game. The mission, in the balance scale condition, was to balance the scale; in the verbal analogy condition, it was to balance the two sides; in the core idea condition, it was to make both sides the same; and in the control group, it was to find the missing number. All students got two trials to find the number that should go in the blank and were given appropriate feedback if they picked an incorrect number.

Screen-shots of each of the four conditions are shown in Figures 3-5. Figure 3 shows the initial set-up for each of the four conditions. For example, in the balance scale condition, the first column in the figure, students saw blocks resting on the pans of a balance scale. The audio asked students if the two sides of the scale were balanced, while accompanying text appeared on the screen (*Figure 3*). Figure 4 shows what happened next for each of the four conditions. Following the example for the balance scale condition, students were asked to balance the scale (column 1 in Figure 4). Correct responses received auditory and dynamic visual feedback, as the scale adjusted to demonstrate balance (column 1 in figure 5). The voice over also verbalized symbolic statements, for example, stating that "3 + 4 + 2 'is the same as' 3 + 6 ", followed by counting of each collection, and "9 is the same as 9…9 equals 9." The equal sign zoomed in corresponding to the moment when the narration said " the same as" or "equals."

Instruction in the verbal analogy applet, was comparable to the balance scale applet, except that the balance scale was absent and was replaced with a set of side-by-side pans (see column 2 in Figure 3) to suggest visual symmetry and the corresponding prompt requested students' to "balance the two sides" (column 2 in Figure 4).

Instruction in the core idea condition was comparable to the verbal analogy condition, except that the two pans were absent, and the prompt asked participants to make the “both sides the same” without any reference to “balancing” the two sides.

Table 2

Instructional Applet Screenshots

Balance Scale	Verbal Analogy	Core Idea	Control

Figure 3. Receiving initial blocks

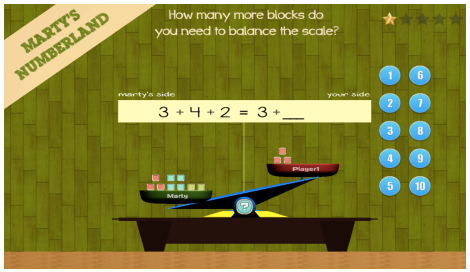	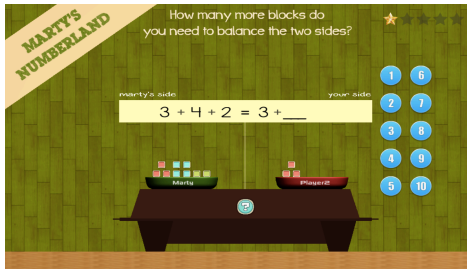	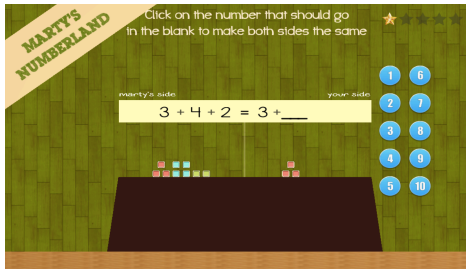	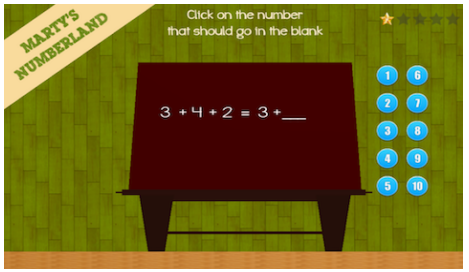

Figure 4. Asking how many more blocks are needed to balance the scale, balance the two sides, achieve sameness, or get a correct answer

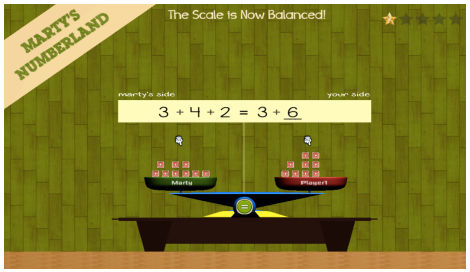	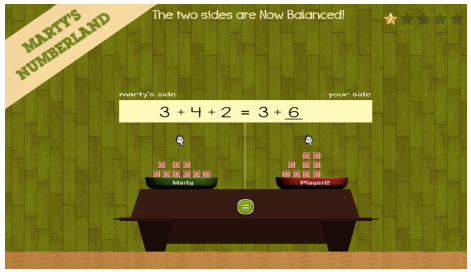	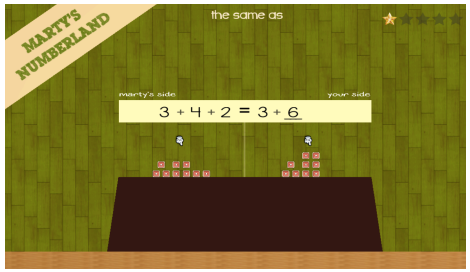	

Figure 5. Receiving final feedback when the number is correct

Note. Instructional conditions show decreasing order of support from left (most) to right (least).

In the control condition, the problems were presented symbolically without any external representation. Students were prompted to find the number that should go in the blank, without any reference to being the same or to balancing. Regardless of their assigned instructional conditions, all participants were exposed to identical problems in the play session (intervention); these are listed in Table 3.

Table 3

Problems Presented During the Intervention: Play session

Problems
$5 = _ + 3$
$2 + 5 = _ + 3$
$3 + 4 + 2 = 3 + _$
$4 + 5 + 2 = _ + 2$
$8 + 3 + 2 = 8 + _$
$2 + 6 + 3 = _ + 3$

Data Sources and Scoring

Students' responses served as data sources. Video recordings also served as data sources for those students with parental approval and student assent.

Scoring the Screening Test

The true-false, addition and subtraction problems were scored either as correct or incorrect.

Scoring of Responses on the Three Components in the Outcome Measures

Equation solving with one unknown. An equivalence problem, referred to as equation with one unknown, was scored as correct if the student provided both a correct number ±1 and gave a relational strategy to the follow-up question to justify why that number should go in the blank. For the practiced problems only, I allowed a response to be counted as correct even if it was off by 1 because many of the children solved the problems by counting and they made frequent counting errors. Moreover, because the focus of this investigation was on promoting a relational understanding of the equal sign (and not fluency in addition), a near miss in computation was acceptable as has also been done by researchers who have studied equal sign understanding in the past (e.g., Perry, 1991). The scores ranged from 0-6 on pretest, posttest, and delayed posttest for practiced problems. Score on transfer items on both of the posttests (immediate and delayed) ranged from 0-3.

Explicit definition of the equal sign. Students' responses to each of the questions were categorized individually based on a system used in prior research (Knuth et al.'s, 2006; McNeil & Alibali, 2005a). In instances where students indicated a relational meaning, even if they also provided an operational meaning, responses were coded as relational. I also categorized responses that indicated a balancing analogy, such as "both sides are balanced" or "it is like the balance" as relational. Interestingly, during the pretest and the posttest some students (n=13) gave a response in which they used the words "same number" to imply a relational meaning but then elaborated that the equal sign goes in the middle, referring to a task typically presented in the first grade. This task provides students with two numbers, an empty box or circle in the middle, and requires students' to identify what symbol (>, <, or =) goes in the empty box in the middle. Some students even elaborated by using the "crocodile mouth," which is a metaphor

used to help students understand the relative inequalities. These students indicated a limited relational view by saying "the same" to imply similarity in numerals (n=n), although it cannot be said for certain how sophisticated their understanding was. During the preliminary analysis, this category was thus identified separately to avoid confounding students' ability to define the equal sign as a relational symbol. For each category, Table 4 provides some examples observed in students' responses that are indicative of their reasoning about the equal sign.

Table 4

Student Response Categories Indicating Explicit Definitions of the Equal Sign

Category	Example
Relational	"it means the same as," "same amount," "like balanced" with a hand gesture
Partial Relational	"it means the same number… like the crocodile, if you have a circle and you have 2 circle 2, you will put equal. If it is 3 circle 2, it is >."
Operational	"it means total," "where the answer goes."
Ambiguous	"it means equals," "like 2 plus 3 *equals* 5."
Other	"I don't know," nodded head –implying no response, or gave an irrelevant response.

Identifying two sides of an equation. Drawn from previous research, this measure had multiple-choice items that provided students with various options, one of which was correct. Responses were coded as either correct or incorrect. The dependent variable was defined as identifying the two sides of the equation correctly in both of the problems that were presented at the pretest, the immediate posttest and the delayed posttest.

The next chapter, reports results on the following four research questions that were examined in the current study:

1) Did the intervention lead to any improvement on the three components of understanding the equal sign at the immediate posttest and was there a significant drop-off after one week on any of these components?
2) Did improvement differ according to the amount of support in the instructional intervention the students received?
3) What were students' affective responses toward the instructional game?
4) Did students' judgments of learning (JOL) map onto their actual learning as assessed by the equation solving performance at the immediate posttest?

CHAPTER 4: DATA ANALYSIS AND RESULTS

Results

Comparability of the Four Groups During the Screening Measure and the Pretest

Students' performances at the screening measure and the pretest are described below.

Screening measure. During the in-class screening test, students scored an average of 3.12 (*SD*=1.33) correct out of 8 true-false problems, 3.17 (*SD*=0.88) correct out of 4 addition problems, and 2.9 (*SD* =0.96) correct out of 4 subtraction problems. Their mean combined proficiency on addition and subtraction questions was 6.07 (*SD*=1.53) correct out of a total of 8 problems. One-way ANOVAs revealed that their performance did not differ across the four instructional conditions in terms of mean scores on true-false [F (3,144) = 0.69, p=0.56], addition [F (3,144) = 2.05, p=.11], subtraction [F (3,144) = 0.88, p=. 45], or addition and subtraction problems combined [F (3,144) =1.03, p=.38)]. Table 5 shows the mean and standard deviation on the screening measure between students' performances in solving any of the problems based on their instructional condition.

Table 5

Mean and Standard Deviation by Instructional Condition on the Screening Measure

Problem Type	Instructional Condition							
	Balance Scale (*n*=40)		Verbal Analogy (*n*=38)		Core Idea (*n*=38)		Control (*n*=32)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
True/False (8 problems)	7.30	2.51	7.50	2.20	7.24	2.17	7.38	2.45
Addition (4 problems)	3.00	1.40	3.32	1.28	2.95	1.27	3.25	1.39
Subtraction (4 problems)	3.00	0.96	3.34	0.75	3.34	0.81	2.97	0.97
Addition + Subtraction	2.73	1.06	2.92	0.78	2.89	1.13	3.09	0.78

Pretest. Students' performances on the three components of understanding the equal sign at the pretest are described next.

Equation solving with one unknown. At the pretest, performance on equation solving was low overall. Students solved 0.22 (*SD* = .65) problems correctly on average (out of 6), with 86% of students (n=127) solving zero problems correctly. Students' performance on equation solving did not differ significantly across the four instructional groups, Welch's ANOVA [$F(3,100) = 0.3$, p=.8)].

Defining the equal sign. At the pretest, the overall percentage of students who gave a relational definition was low (8.1%), with the majority of students giving an operational definition (73%). Additionally, about 5.4% of students gave an ambiguous definition; while about 4.7% either stated that they did not know what the symbol meant or gave a response that was not relevant (these last two types of response have been combined under the category of "other"). Approximately, 8.8 % gave a partial relational definition. Students' ability (or lack of ability) to define the equal sign relationally did not differ according to their instructional condition during the pretest, Fisher's Exact (p= .90).

Identifying the two sides of an equation. At the pretest, only 10% of students identified correctly the two sides in both of the problems that were asked. The percentage of these students did not differ across instructional conditions, Fisher's Exact (p=0.1).

Table 6 provides students' pretest performances by instructional condition on the three components of understanding the equal sign.

Summary of results for the screening measure and pretest. Taken together, these results indicate that students' performances did not differ significantly among the four instructional conditions on any type of problems given during the screening measure or on any of

the three components of understanding the equal sign measured during the pretest. Thus, the four groups were comparable prior to the intervention.

Table 6

Performance on Each of the Three Components of Understanding the Equal Sign by Instructional Condition

Component	Performance Measure	Balance Scale	Verbal Analogy	Core Idea	Control
Equation solving with one unknown	M (SD)	0.20 (0.52)	0.18 (0.51)	0.21 (0.74)	0.31 (0.82)
Defining the equal sign	% who defined it relationally	15.0 %	15.8 %	13.2 %	25.0 %
Identifying the two sides	% who got both correct	2.5 %	7.9 %	13.2 %	18.8 %

The results from the pretest also indicate that, overall, students did not understand the equal sign in relational terms, as assessed by their poor performance on all three components of understanding the equal sign.

Evidence of Intervention Effectiveness

To determine whether the interventions were effective, this study evaluated student performance on the three components of understanding the equal sign at an immediate posttest and again after one week. The one-week delayed posttest was used to determine whether there was any drop-off in any initial gains students had made, as measured by the immediate posttest.

Immediate posttest. Students' performances on the three components of understanding the equal sign at the posttest are described next.

Equation solving with one unknown. Students performed better at the posttest than at the pretest, with students solving an average of 2.51 (*SD*=2.54) out of the 6 *practice* problems correctly (i.e., problems in the same format as those practiced during the intervention phase) and

an average of 0.60 (*SD*=.84) out of the 3 *transfer* problems correctly (i.e., problems in formats that differed from those provided during the pretest or intervention phase). Additionally, students' performance on transfer problems correlated significantly and positively with their performance on practice problems (*rho*=.79, *p*<.001). Thus, remaining analyses combined performance on practice and transfer items.

Defining the equal sign. The percentage of students who defined the equal sign relationally increased from 8.1% at the pretest to 28.4% during the immediate posttest, while the percentage of students who gave an operational definition decreased from 73% to about 34%. Additionally, there was an increase in the percentage of students who reported an ambiguous definition as well an increase in the percentage of students who gave no response or reported not knowing what the equal sign means. Figure 6 compares the percentage of students whose responses fell into each definitional category at the pretest and posttest.

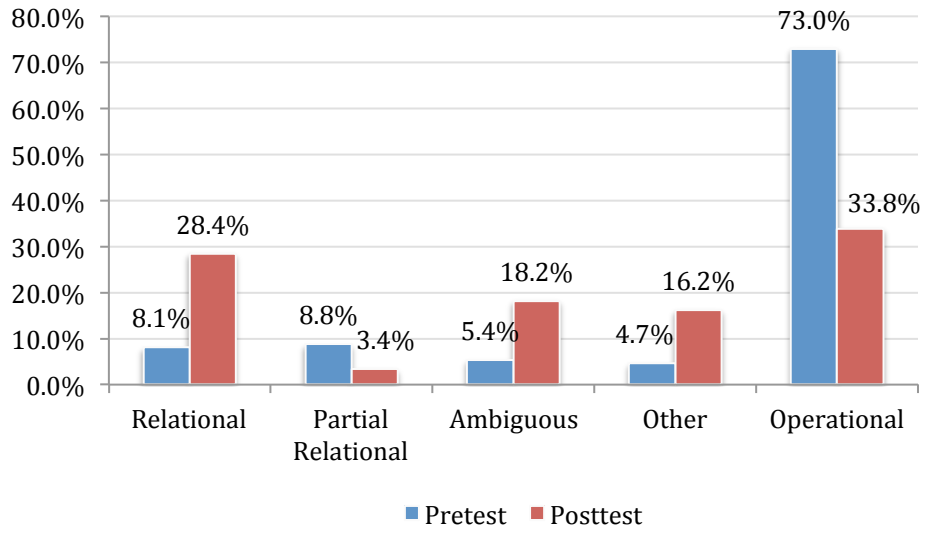

Figure 6. Percentage of students in each of the equal sign definition categories at the pretest and posttest (*N*=148)

***Identifying the two sides of an equation*.** The percentage of students who identified the two sides of an equation relationally increased significantly from the pretest (10%) to the immediate posttest (28%), McNemar's χ^2 [(1, 148) =12.595, p<.001].

Taken together, these results indicate that students' performance improved at the immediate posttest across the three components of understanding the equal sign (solving equations correctly, defining the equal sign relationally, and identifying the two sides of an equation).

Delayed Posttest. To determine whether there was a significant drop-off in performance one week after intervention, non-parametric tests were conducted. These tests measured the differences between students' performances on the three outcome measures on the immediate posttest and the delayed posttest.

Equation solving with one unknown. Due to violating assumptions for a parametric test, Wilcoxon rank-tests within each group were considered and reveal that students' proficiency in solving an equation with one unknown correctly did not differ between the immediate posttest and the delayed posttest; there was no significant drop-off after the immediate posttest (all p>.05). Table A1 (see Appendix A) provides test statistics and p values for each of the problem types, considered by instructional condition.

***Defining the equal sign*.** A McNemar's test indicated that there was no significant difference [χ^2 (1, 148) = .03, p =.56] between the immediate and delayed posttests (when corrected for contiguity) in whether students gave a relational definition or not.

Identifying the two sides of an equation. A McNemar's test indicated that there was no significant difference [χ^2(1, 148) = 0, p =1] between the immediate and delayed posttests (when corrected for contiguity) in whether students identified the two sides of the equations correctly.

Taken together, these results indicate that the overall effects of the intervention did not differ significantly from the time of the immediate posttest to that of the delayed posttest. Because there was no significant drop-off in students' performance between the immediate and the delayed posttests, in the remainder of this chapter, I ignore the delayed posttest and present analyses that use only the immediate posttest to examine the differential impacts of the four instructional conditions.

Examining the Effects of Interventions on Performance

Analyses were conducted with one independent variable, intervention, with four levels (balance scale, verbal analogy, core idea, and control). To assess whether there was any improvement in students' performance in solving equations, the dependent variable was a composite score (i.e., gain from pretest to posttest + transfer). As mentioned previously, a correct score was based on accuracy of the response as well as appropriate justification of arriving at that response (e.g., relational strategy). To assess differences in students' performance on defining the equal sign, the dependent variable was the definition that students gave (relational or partial relational = 1; anything else = 0). To assess students' performance on identifying the two sides of an equation, the dependent variable was a combined score for students' responses to the posttest questions (correct on both = 1; incorrect on one or both questions = 0).

Improvement on equation solving. Although, in general, students improved after the intervention, they did not do so equally across all four interventions. Results indicate that students in the verbal analogy condition benefitted most. Students in the core idea condition also benefitted, but not as much as students in the verbal analogy condition and were next. Students in the balance scale condition benefitted the next. These results apply, irrespective of the indicator (e.g., transfer) used to assess performance at the posttest. Table 7 provides detailed information

on students' performance on equation solving problems during the posttest, given their instructional condition.

Table 7

Posttest Performance on Equation Solving Problems by Instructional Condition

Performance	Instructional Condition							
	Balance Scale (*n*=40)		Verbal Analogy (*n*=38)		Core Idea (*n*=38)		Control (*n*=32)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Practice problems	2.50	2.44	3.71	2.50	2.76	2.63	0.78	1.62
Transfer problems	0.50	0.78	0.84	0.89	0.82	0.95	0.19	0.54
Gain on practice problems	2.30	2.42	3.53	2.52	2.55	2.56	0.47	1.22
Improvement*	2.80	3.05	4.37	3.20	3.37	3.40	0.66	1.56

Note. *Composite score on practice gains and transfer

Analysis of variance on improvement scores. During the posttest, the assumption of normality within the instructional groups was violated (Shapiro-Wilk, p<.000), for a one-way ANCOVA on the improvement scores. A one-way Welch's ANOVA was used, because the Levene's test for homogeneity of variances revealed unequal variances (p= .000). Analyses revealed statistically significant differences in gain scores across the instructional groups, Welch's F [(3, 78.13) = 17.577, p = .000]. A Games-Howell post-hoc comparison (at $p < .05$) confirmed that students in the three experimental conditions gained significantly more than students in the control condition: verbal analogy (M = 4.37, SD 3.19), core idea (M = 3.37, SD = 3.39), balance scale (M = 2.80, SD = 3.05), and control (M = .66, SD = 1.55).

Categorizing levels of learning on solving equations. The frequency distribution of students' performance on equation solving was skewed (i.e., not normally distributed), which is typical in studies of children's performance on such problems (McNeil, 2007). This is usually

handled by categorizing students' responses binomially in terms of whether or not students solved at least one of the problems correctly.

Although categorizing responses binomially allows researchers to assess whether students learned or not, it does not allow us to evaluate gradations of learning. Therefore, given the nature of the data and the precedent set by previous research, I examined both logical and empirical ways to classify outcomes into levels of learning. The classification system I developed accounted for gain scores (i.e., even though students were comparable across pretest conditions, these learning levels accounted for any slight variability between individuals) and distinguished between levels of learning by considering learning on an ordinal scale rather than a binary scale (see Agresti, 1996).

To aid in identifying levels of learning, I classified students' performance into three levels: a) *none*, demonstrating no gain on the practiced problems nor on any of the transfer problems (n = 61); b) *adequate*, demonstrating a *minimum gain of at least five* on the composite scale, with, in particular, a *gain of at least four* on practice problems and *at least one* on a transfer problem (n = 45); and c) *emerging*, demonstrating some gain between none and adequate, on *combined gain of four* on practice and transfer problems. I chose a gain of four problems as the cut off, because, logically, students getting two-thirds of the problems correct (after accounting for pretest performance) rules out the likelihood that they guessed or that their correct responses were due to chance. Empirically, the vast majority of students who had a gain of *at least four* at the posttest also answered at least one transfer problem correctly (n=45 out of 53). This was not the case for some students who solved *up to five practice problems alone* (n=8), indicating that their performance was, emergent. These criteria allowed for the examination of nuanced patterns of learning with respect to the different instructional conditions.

Analyzing intervention effects on levels of learning. Table 8 shows the percentage of participants in each of the four instructional conditions who exhibited different levels of learning. A Fisher's Exact test (p=.001) revealed that a student's level of learning was related to the amount of support he or she received (i.e., the type of instructional condition). To pinpoint which conditions led to different levels of learning, I analyzed the data with ordered logistic regression.

Table 8

Percentage of Participants in each Level of Learning by Condition at the Posttest

Levels of Learning	Instructional Condition			
	Balance Scale (n=40)	Verbal Analogy (n=38)	Core Idea (n=38)	Control (n=32)
None	45%	16%	37%	63%
Emerging	28%	37%	24%	34%
Adequate	28%	47%	40%	3%

Note. None is the same as no learning

To examine whether the odds of reaching a higher level of learning differed significantly by condition and to identify influences of demographics on the levels of learning, I conducted an ordered (also known as ordinal) logistic regression with three levels of the outcome variable (i.e., none, emerging, and adequate) and the intervention condition as a factor with four levels (i.e., balance scale, verbal analogy, core idea, and control). To identify demographic influences—age, gender, socio-economic status (SES, based on eligibility for free or reduced lunch, with 1= eligible, low SES and 0 = not eligible, high SES), and English Language Learner (ELL) status (0 = not ELL; 1 = ELL) were first included in the model as predictors. Gender, age, and ELL status were not significant predictors in the base model. Even though age was not significant, because it has been a significant predictor in solving equivalence problems in previous research (see

especially McNeil, 2007), I allowed age to remain in the model. Note that none of the results changed when age was excluded from the model.

Ordered logistic regression results with the control condition as the reference. After removing gender and ELL status, and keeping the other predictors in the model, I again performed an ordered logistic regression, with the control condition as the reference level. In other words, in this model, the control condition and higher SES were coded as zero. The model estimated a log of the odds of students being in a higher learning level (or levels) relative to the lower learning levels based on each of the predictor variables, using the control condition and higher SES as the reference.

Table 9 provides detailed information on the results from the ordered logistic regression model, including coefficient estimates (β), standard errors (se), t-values (t), and the odds ratio (OR) with corresponding confidence intervals for each of the predictor variables used. The results reveal that students in the balance scale condition, who received the most support, had the lowest odds ratio and the students in the verbal analogy condition, who received the second-most support, had the highest odds ratio. Specifically, for students receiving the verbal analogy intervention, the odds of being classified as achieving *adequate* versus *emergent and no* learning levels combined were 8.32 times greater, when compared to being in the control condition with all other variables in the model held constant ($p<.001$). For students in the core idea condition, the odds of achieving *adequate* versus only achieving *emergent understanding or not* learning were 4.07 times greater when compared to being in the control condition, when all other variables in the model were held constant ($p=.003$). And, for students in the balance scale condition, the odds of achieving *adequate* versus only achieving *emergent understanding or not*

learning were 2.69 times greater when compared to being in the Control Condition, when all other variables in the model were held constant (p=.04).

Table 9

Ordered Logistic Analysis on Levels of Learning at the Posttest, with Control as Reference

Predictor Variable	β	se	t	*p*	Odds Ratio	95% CI	
						2.5	97.5
Balance Scale	0.99	0.47	2.08	0.04*	2.69	1.07	7.01
Verbal Analogy	2.11	0.48	4.33	0.00*	8.32	3.27	22.38
Core Idea	1.40	0.48	2.92	0.00*	4.07	1.61	10.68
Eligible For Free Lunch	-1.18	0.38	-3.08	0.00*	0.30	0.14	0.64
Age	0.13	0.23	0.53	0.60	1.14	0.71	1.82
AIC ** =	305.07						

Note. ** Akainine value for the model; * p<.05.

The control group and high SES (non-eligibility for reduced lunch) group are references in the model.

Moreover, for students with a higher SES, the odds of being classified as having achieved an *adequate learning* level versus either *emerging learning* or *no learning* level were 3.33 times greater (using an inverted odds ratio, since the coefficients were negative for this predictor) than for students with a lower SES, when other variables remained constant (p=.003). For a one-year increase in age, the odds of students being classified as having achieved an *adequate* level of learning versus *emerging or no* learning were only 1.14 times greater, (p=.60).

Thus, the three interventions with support (balance scale, verbal analogy, and core idea) were more likely to result in higher learning than the control condition. Students' with higher SES also had greater odds of improving over students in lower SES.

Next, to examine if being in the condition that provided the most support, the balance scale condition, led to differential learning, based on a priori hypothesis, I reanalyzed the data by

considering the balance scale group as the reference and comparing it to the other two treatment conditions (verbal analogy and core idea), which I refer to as "some support".

Ordered logistic regression results with the balance scale condition as the reference. This model estimates a log of the odds of students achieving a higher level of learning relative to a lower level. Table 10 provides detailed information on results from the ordered logistic regression, including coefficient estimates (β) and the odds ratio with corresponding confidence intervals for each of the predictor variables used.

Table 10

Ordered Logistic Analysis on Levels of Learning with Balance Scale Group as the Reference

Predictor Variable		β	se	t	*p*	Odds Ratio	95% CI	
							2.5	97.5
Some support		0.79	0.38	2.09	0.03*	2.21	1.05	4.70
Control		-0.97	0.47	-2.04	0.04*	0.38	0.15	0.95
Eligible For Free Lunch		-1.12	0.38	-2.98	0.00*	0.32	0.15	0.67
Age		0.20	0.23	0.86	0.39	1.22	0.77	1.94
AIC**	=	305.63						

Note. ** Akainine value for the model; * p<.05.

The groups with the most support and higher SES were references in this model.

The results revealed that students in the verbal analogy and core idea conditions were 2.2 times more likely than students in the balance scale condition to show *adequate* versus *emerging or no* learning, when all other variables were held constant (p = .03). Similarly, the odds of students being at an *adequate or emerging* learning level when compared to *not learning* was 2.2 times greater for students in the verbal analogy and core idea conditions when compared to students in the balance scale condition, with all other variables remaining constant (p=.03).

Moreover, for students with a higher SES, the odds of being classified as being at an *adequate* versus *emerging and not learning* level were 3.1 times (using an inverted odds ratio, since the coefficients were negative for this predictor) than for students from a lower SES (p=.003). For a one-year increase in age, the odds of students being classified as *adequate* versus *emerging and not learning* was only 1.2 times greater (p=.39.), which was not a significant advantage.

Taken together, these results indicate that students who received some support, as compared to no support, significantly increased their odds of showing adequate or some learning on the equation-solving component of understanding the equal sign. Additionally, when considering amounts of support, the verbal analogy and core idea interventions combined were more predictive of higher levels of learning when compared to the balance scale intervention.

Students' performance on defining the equal sign. A Fisher's Exact test ($p<.001$) indicated that students' proficiency in defining the equal sign relationally differed significantly across the four instructional conditions. Following this test, I used binomial logistic regression to predict the log of the odds of students giving a relational definition (see Agresti, 1996). For this analysis, student responses that were either relational or partial relational definitions were coded as relational, since results did not differ when partial relational definitions were considered as a separate category. Instructional conditions (with control as the reference group), SES, and age were entered as predictors in the model. The overall model estimated that the odds of students giving a relational definition were significantly higher in any of the other conditions when compared to the control (*Wald* χ^2= 12.7, p=.005). Table 11 shows detailed information, including the estimate coefficients (β) for each of the predictor variables and their respective odds ratio. Specifically, when compared to the control condition, the log of the odds of students giving a

relational definition were significantly increased: by 16.24 [3.97, 111.64] on receiving the verbal analogy intervention, by 7.62 [CI = 1.83, 52.43] on receiving the core idea intervention, and by 6.48 [CI = 1.54, 44.91] on receiving the balance scale intervention.

For SES, the inverted odds ratio indicated that the log of the odds of students giving a relational definition (versus not) was 2.85 times higher for those with a higher SES than for those with a lower SES (p= .02) CI [0.15, 0.86]. For age, the log of the odds of students giving a relational definition (versus not) increased by 1.26 times for one year increase in age; this was not significant, CI [0.72, 2.20].

Table 11

Binomial Logistic Regression Results on Relational Definition of the Equal Sign

Predictor Variable	β	se	z	*p*	Odds Ratio	95% CI	
						2.5	97.5
Balance Scale Condition	1.86	0.82	2.28	0.02*	6.48	1.54	44.91
Verbal Analogy Condition	2.79	0.81	3.43	0.00*	16.24	3.97	111.64
Core Idea Condition	2.03	0.82	2.49	0.01*	7.62	1.83	52.43
Eligible For Free Lunch	-1.02	0.45	-2.29	0.02*	0.36	0.15	0.86
Age	0.23	0.28	0.80	0.42	1.26	0.72	2.20

Note. * p<.05.

Using Wald's test, I conducted follow-up comparisons to find whether the coefficients for each of the instructional condition pairs were equal (e.g., whether the coefficient for the balance scale instructional condition was equal to coefficient for the verbal analogy condition). This analysis found a marginally significant difference between the coefficients of the verbal analogy and balance scale conditions in the model that used the control condition as the reference (*Wald* χ^2= 3.6, p=.05). No other comparisons were significant (all p>.05).

The results indicate that during the posttest students in the verbal analogy condition were more likely to give a relational definition (than not) when compared to those in the control condition. They were marginally more likely to give a relational definition when compared to those in the balance scale condition, with other factors remaining constant.

Students' performance on identifying the two sides of an equation. At the posttest, the number of students correctly indicating the two sides of an equation did not differ significantly by instructional condition [$\chi^2(3,148) = 2.60$, $p = .46$]; hence, no follow-up tests were considered.

Students' Reponses on the One-Time Survey Administered at the Posttest

All 148 students completed a survey that had a total of nine questions, with six questions on a 4-point Likert scale (see Appendix D for survey questions). The affective survey was administered to examine whether participants in certain groups had a higher response towards liking the game, finding math to be fun in the game, and wanting to get more difficult problems in the game they had just played based on the condition they received. Another objective of the survey was to examine whether students' judgments of whether the game helped them in solving math problems actually reflected their learning at the immediate posttest as assessed by their performance on equation solving problems.

Affective responses to instruction in the virtual environment. Figure 7 depicts students' mean scores on various survey questions, along with the percentage of students in each response category on a 4-point Likert-scale. Overall, most students gave a high rating to the survey questions, with the exception of a question that asked them whether they would like to get more difficult math problems. In the figure, means are reported only to indicate that there were

more students choosing “agree” (3) or “absolutely agree” (4) than either of the two options for disagreement. Overall, the questions did not distinguish between the groups.

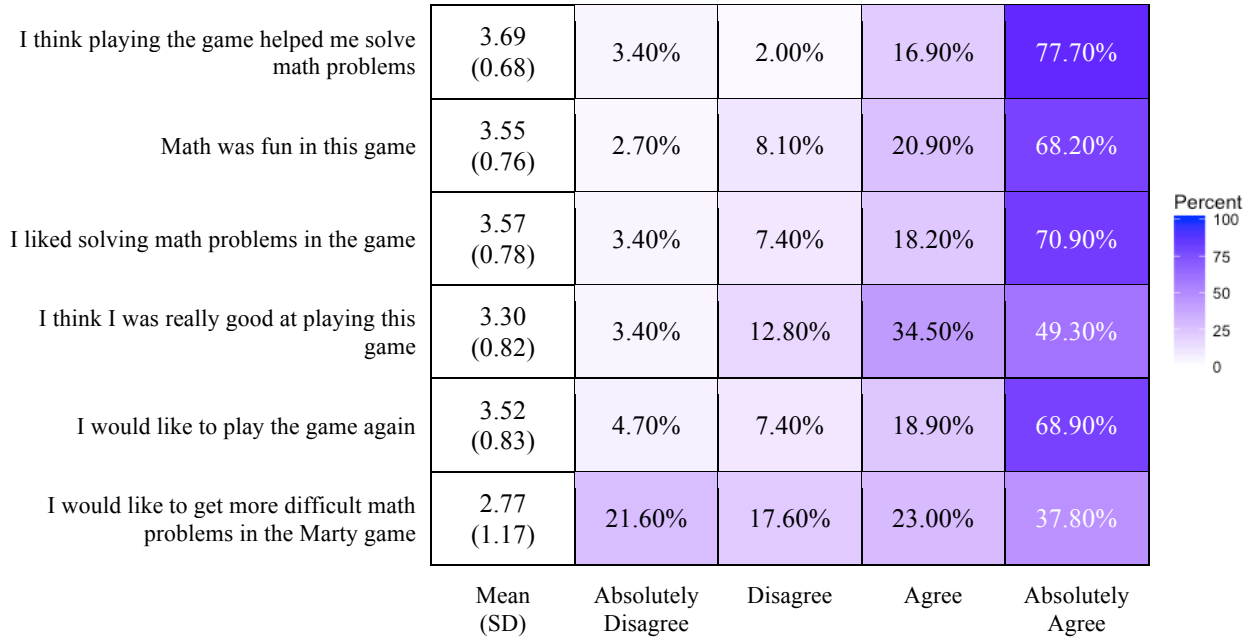

Figure 7. Heat map with mean and standard deviations of responses to the Likert-scale survey questions; darker color depicts higher percentages.

Time spent playing on a computer in a week. To take into account any differences in students’ frequency of computer use, one question asked, “How often do you play on a computer or laptop at home in a week?” Analyses revealed no significant differences between the four instructional conditions [χ^2 (9,148) = 8.41, p =.49].

Students’ report on the likability of solving math problems in the game. Figure 8 provides density plots of students’ ratings on the six Likert-scale questions. Although the majority of the students reported that they found math to be fun in the game, the students who were assigned to the verbal analogy condition, on average, gave the highest ratings to questions that asked whether they wanted to play the game again and found math to be fun. The average

ratings of students in the verbal analogy condition were almost the same as the average ratings of students in the balance scale condition on questions that asked if students liked solving math problems in the game; would like to get harder problems in the game; and thought the game helped them to solve math problems. Students' ratings on whether they would like to get more difficult math problems in the game differentiated the balance scale and the verbal analogy conditions from the core idea and control conditions.

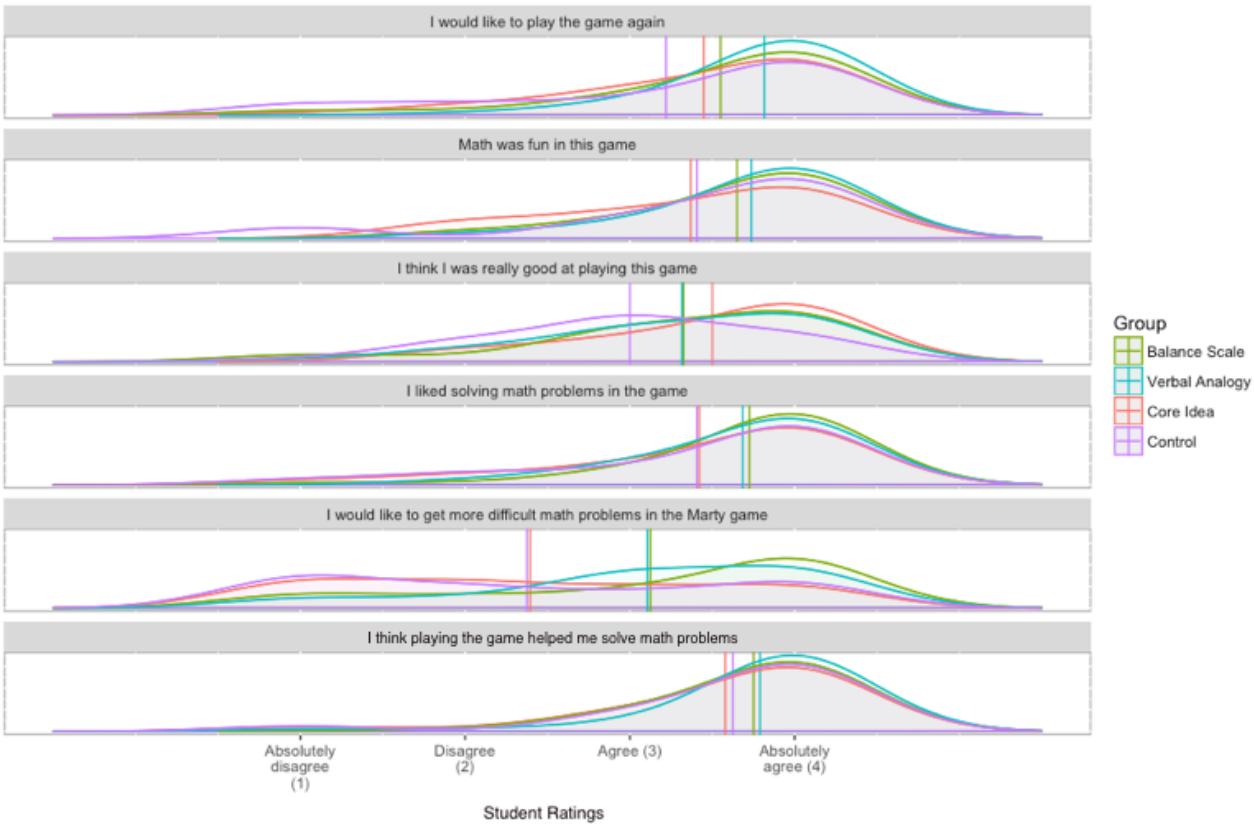

Figure 8 Density plot of different responses to survey questions by instructional condition

To examine if there was a significant difference in students' responses across the four conditions, I ran an ANOVA on student responses as the outcome variable (1 to 4) for each of the questions, with instructional condition as the predictor. There were no significant differences across the groups on students' reports on whether they found math to be fun in the game [$F(3, 144) = 2.15$, $p = .09$]; whether they thought they were good at playing the game [$F(3, 144) =$

2.25, p = .08]; or whether they liked solving math problems in the game, [F (3, 144) = 1.17, p = .15].

A Fisher's exact test was conducted to examine whether there were difference in how much students reported liking to solve math problems in the intervention (game); it revealed non-significant differences (p=.59) across the four instructional groups. These results were somewhat surprising, and indicated that students in the control condition reportedly liked the game as much as students in the other conditions despite the fact that they received almost no support when compared to the other three conditions, although other features (e.g., animated character, were kept similar).

Students' judgments of learning. To examine whether there was a significant difference across the four conditions, on students' self-judgments on their learning, a Fisher's Exact test was used and was not significant (p=.43). That is, there were no differences across the conditions on students' reports on their self-judgments of learning.

To examine if students' self-judgments of learning matched their performance on equation-solving problems, I converted the response to the survey question "I think playing the game helped me with solving math problems" from a 4-point to a 2-point Likert scale to consider extreme categories. Students' performance on solving equations was categorized as achieving either at least an emergent (i.e., emergent or adequate) or not (i.e., did not show any learning) level of learning. This 2 X 2 Fisher's Exact test was not significant (p=1), indicating that students' judgments of their learning did not correspond with their actual learning on the posttest.

CHAPTER 5: DISCUSSION AND CONCLUSION

Overview

There has been decades of research on elementary school children's understandings of the equal sign as well as recommendations to use the pan balance as a concrete tool to teach this mathematical concept (NCTM, 2000; CCSSM, 2012). Yet the current study was the first to examine, in a randomized controlled experiment, what features (i.e., what amounts of support) associated with the balance scale manipulative impact the development of a relational understanding of the equal sign in children. The absence of previous studies that specifically address the effects of the balance scale on children's learning may be partly due to the intuitive assumption that this manipulative works. This is reasonable because, at the visual and mechanics level, this manipulative demonstrates the idea that the equal sign is meant to convey in arithmetic: namely, equivalence. These characteristics make the balance scale appear to be a well suited manipulative that offers optimal external support for teaching children about the relational meaning denoted by the equal sign.

The findings of the current study question the intuitive appeal of providing apparently optimal support for teaching and learning in mathematics, especially when using this manipulative. In particular, offering young students an opportunity to interact with the vifrtual balance scale was not as effective as offering young students a more stripped-down version (i.e., a verbal analogy of the balance scale or presentation of the core idea embedded in the balance scale). Therefore, this study adds support to theories that propose a distraction account of the use of rich manipulatives in the teaching and learning of abstract concepts in mathematics. Specifically, this study provides evidence that while students' relational understandings of the equal sign can be supported with the virtual balance scale manipulative in a brief instructional

episode, there are other features associated with the manipulative that have advantages over the use of the fully enacted manipulative in promoting students' learning about the relational meaning of the equal sign.

In the remainder of this chapter, I discuss some of the more intriguing findings: 1) young children can be taught the relational meaning of the equal sign in a brief intervention; 2) results may have been enhanced because the instruction was delivered on a computer; 3) results may also have been enhanced because of the use of the not-equal and equal signs; and, most significantly 4) the intervention providing the most support did not lead to the most impressive learning outcomes.

After the discussion of these findings, I discuss implications for teaching and learning, limitations and future directions, and end with conclusions.

Discussion of the Intriguing Findings

Young Children Can be Taught the Relational Meaning of the Equal Sign

Results indicated that participants in the targeted age group could perceive the equal sign relationally, as demonstrated by their performance on all of measures of understanding the equal sign used in the present study. Previous research McNeil's previous research (2005, 2007) found that children between 7 and 9 years of age tend to have an operational view of the equal sign; McNeil reckoned that children's understanding of the relational meaning of the equal sign was hampered by their entrenched experiences of using the equal sign operationally. However, the children in McNeil's studies were found to be less responsive to instruction than the children in the present study. This makes the findings from this study particularly intriguing because the intervention was relatively brief (~30 minutes) and primarily self-regulated, given the constraints and affordances of the multimedia environment. The findings from the current study thus support

findings from previous studies (Matthews, & Rittle-Johnson, 2009; McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011; Perry, 1991; Rittle-Johnson & Alibali, 1999) that employed traditional one-on-one instructor-guided interventions and demonstrated that small changes in instruction can help children see the equals sign relationally. However, the findings in the current study differ from those that found age to be predictive of whether students would benefit from instruction or not (e.g., McNeil, 2005, 2007).

Considering the Effects of Instruction Delivered on a Computer

A distinguishing feature of the current intervention was that it was delivered in a multimedia environment. According to multimedia theory of learning (Mayer, 2005c; Clark & Mayer, 2011), information that is presented in both auditory and visual forms in close proximity supports learning by reducing the cognitive load in working memory. Thus, introducing the visual symbol (=) and the audio message ("the same as") in sync in the interventions likely contributed to students' improved performance. Specifically, with respect to learning in mathematics using manipulatives, some research has indicated that there are two particular advantages to computer renditions of physical manipulatives: a) the flexibility available in maneuvering the virtual manipulative, which emerges from the fact that "clumsiness" is avoided in the case of virtual manipulatives, and b) the connection between the symbols and the actions (on the virtual manipulative or image) because the feedback is immediate and can be directly linked to the object of study (Clements, Battista, & Sarama, 2001). Moreover, according to trainers' observations, overall, students were excited, liked to play on the computer and seemed to be engaged interacting with the content delivered during the game-like instructional phase. Students' responses to survey questions also revealed that they liked playing the game and reported that they would have liked to play it again. Thus, the overall effectiveness of this brief

intervention without significant differences observed in age could be in part due to presenting the instruction in a multimedia format on the computer.

The Advantage of Introducing the Equals Sign With a Not-Equals Sign

The observed overall effectiveness of the interventions—and age not being a significant predictor—could in part be attributed to the fact that students were introduced to the equal sign along with the not-equal sign. For example, during the intervention, when the blocks did not share the same amount on both the sides, students were introduced to a not-equal sign (≠), along with a voice-over stating "this side (the left side) is not the same as this side (the right side)". In fact, Hattikudur & Alibali (2010) found that when third- and fourth-graders were introduced to the equal sign along with the two inequality symbols of greater than and less than, they demonstrated higher gains as compared to those who received instruction only on the equal sign. Thus, introducing the equal sign as a symbol of equality along with introducing the not-equal sign might have contributed to learning in the current study. Future research on comparing this aspect of the intervention could strengthen the claim that introducing the equal sign in conjunction with the not-equal sign is effective for breaking students' problematic reliance on the equal sign as an operational symbol.

Interestingly, during the posttest, in addition to an increase in the percentage of students who gave a relational definition, there was an increase in the percentage of students who reported an ambiguous definition or reported that they did not know how to define the symbol. This suggests that for the students who did not fully embrace the relational view of the equal sign, there was some disruption in their views of the equal sign as an operational symbol after the brief intervention. These students were emergent because they abandoned the operational definitions of the equal sign but did not quite accept the relational view.

The Most Support May Not Be the Best: Amount of Support Matters Within a Brief Instruction

One of the main questions this study examined was whether the instruction using the more support or whether using limited features of the manipulative would be more conducive to learning. Analyses of improvement on the posttest revealed that participants who received at least some amount of external support on all or some of the features associated with the balance scale performed significantly better than those who only received feedback in terms of correctness or incorrectness of the solution (i.e., the control condition). This finding was not surprising, as one would expect students' understanding of an abstract concept, such as equivalence denoted by the equal sign, to be facilitated by some concrete representation and contexts that are non-symbolic (Sherman & Bisanz, 2009). Moreover, previous work has indicated that explicit instruction helps students learn about the equal sign (e.g., Fyfe et al., 2014; McNeil & Alibali, 2005; Molina & Ambrose, 2006; Perry, 1991).

However, given the age of participants and their developmental stage—concrete operations (Piaget, 1970)—one might expect that students would learn more when their ideas are supported by the instruction that incorporates manipulatives, since a complete "manipulative enables students to connect abstract mathematical concepts to real objects" (Kennedy & Tipps, 1994, p. 71). The appeal of manipulatives in mathematics education has been long observed, and a complete manipulative seems to provide structure for teaching (and arguably learning). Thus, it may be expected that instruction that facilitates equivalence understanding using the intervention of the balance scale would be more beneficial than any of its stripped-down versions. Surprisingly, this was not the case in the current study. Findings revealed that all other features being similar, receiving ***less than*** the most support, or in this case receiving instruction *without*

the presence of a virtually dynamic balance scale, led more students to solve more equivalence problems accurately as compared to the students who received instruction using the most support (i.e., the balance scale condition). These findings, though unexpected, indicate that providing the most support with the dynamic manipulative was less disruptive and less productive in altering children's operational views of the equal sign. Lesser amounts of support were more effective, with the condition that provided instruction using the analogy of balancing the two sides as the most disruptive. This was true for students' performances on solving equations correctly as well as providing relational definitions of the equal sign. I now consider several reasons that less support may have led to greater learning.

The disadvantage of the most support. Other researchers (e.g., Brown, McNeil, & Glenberg, 2009; Kaminski, Sloutsky, & Heckler, 2009; Petersen & McNeil, 2013; Uttal et al., 2009) have made the claim that manipulatives that are perceptually rich may add to cognitive load, may constrain learning, or draw students' attention away from the intended meaning of the manipulative. This study adds to the evidence, supporting theories that predict the negative impact that some manipulatives can have on students' learning (Martin, 2009; Uttal, et al. 2009). For instance, Martin (2009) provides a theoretical framework for understanding why realistic manipulatives can hinder the development of generalizable mathematical ideas. According to the framework of physically distributed learning, children come to develop general structures for solving problems without the materials by repeating "action-interpretation structures" they form after working on the materials (p.141). This generalization is crucial for learning. Martin warns that the concrete materials may sometimes do too much of the work for the learner and thus may constrain the abstraction of actions-interpretation structures, which causes the materials to hinder both learning and transfer. Because the balance scale by its very nature incorporates the idea of

equivalence, it may not have allowed the participants to disassociate the meaning of "balance" from the mechanics and image of the scale.

The theory of physically distributed learning and perspectives on why perceptually rich manipulatives can be distracting (e.g., Martin, 2009; Kamiski, Sloustsky & Heckler; McNeil & Uttal, 2009) also provide the basis for interpreting why the analogy and the core idea interventions led more participants to achieve higher gains than actually seeing the balance scale in action. With the presence of the balance scale, the visual was structured and complete; therefore, it was so tightly tied to the idea of equivalence that it may not have offered enough opportunity for students' reflections. This kind of reflection may be necessary to transfer from the context of the balance scale to arithmetic; or in Martin's words, to form action-interpretation structures. In the analogy condition, some level of dissociation between the idea of balance and equivalence was already present, possibly allowing for the disruption of students' faulty patterns, as they had to think mentally about the idea of balance without relying on an image of a balance scale. In the core idea intervention, again, the students did not have a perceptually rich image. In other words, both the analogy and core idea conditions provided the idea of equivalence without tying it to the image and dynamic movement of the balance scale; this may have allowed students to reflect on the idea of what it means to balance or to simply think about making the two sides the same. In essence, the problem with the balance scale was that the students did not have to engage in the work themselves; by only getting the analogy or core idea, they had to come up with a way to make sense of this for themselves.

A second rationale to explain why the less supportive interventions were more beneficial is based on the idea that paying full attention to the most supportive intervention was demanding,

thus draining cognitive resources necessary for processing the ideas being demonstrated in the intervention (Chandler & Sweller, 1991; Kaminski, Sloutsky, & Heckler, 2009).

A third, and quite different, theoretical perspective, based on the embodied nature of learning (Lakoff & Nunez, 2001; Lindgren & Moshell, 2011), would suggest that providing the most visual support might have been less beneficial because freeing memory space (i.e., providing the most support) does not automatically result in learning. On the contrary, the more that students can engage their mind *and body*, the more likely the activity will result in greater and deeper learning. In light of this recent work, more support would mean involving sensory and kinesthetic faculties of the learner within a multimedia environment. Proponents of embodied learning argue that engaging the audio, visual, and kinesthetic senses can lead to a deeper learning by immersing various faculties into meaning making. In particular, abstract concepts that are hard to grasp through direct instruction using audio-visual support may be especially responsive to learning through embodied approaches. However, whether this perspective impacts students' learning of mathematical equivalence is an open question. Although some work has found positive effects when students actually balance a scale (i.e., different than the pan balance scale used in this study [Perry, Berch, & Singleton, 1995]), it would be interesting to apply an embodied learning perspective to support students' understanding of the equal sign. The open question is whether using the balance analogy while engaging more faculties can indeed lead to young elementary school students to abandon their operational patterns and successfully move to adopt a relational understanding of the equal sign. Positive results might add to the intervention approaches towards developing mastery in equivalence problems and a deeper understanding of this essential and omnipresent symbol, and might also indicate the pedagogical effectiveness of the balance scale as an immersive rather

than a visual model. In fact, some classroom-based work (Mann, 2004) highlights that students were more accepting of equivalence problems when teachers had them model problems using their stretched arms to represent a see-saw (i.e., another model of balance) during the academic year (Mann, 2004).

Implications For Teaching and Designing Instruction

A serious implication of this project is that although manipulatives may be useful, educators need to evaluate their pedagogical effectiveness. For teachers, this means that at best, providing seemingly helpful manipulatives may not allow students to reflect and meet their complete potential and at worse may hinder learning or give the illusion that they are learning. Recent work has indicated that using a concreteness-fading approach (Fyfe et al., 2014), i.e., one in which students begin by working with concrete objects then the objects are gradually replaced by symbols, is effective in allowing students to bridge the gap from the concrete to the abstract.

The current study adds nuance to this important finding. It may be more effective to *not* begin with a completely embellished manipulative or representation that demonstrates the concept and instead provide just-enough support, leaving students to take up much of the cognitive work on their own. Uttal et al. (2009) and others (e.g., Kaminsky & Sloutsky, 2013) have made strong arguments for bare-boned representations to remove distraction. The work presented here supports this but goes farther by presenting instruction on the computer, which further limited students' abilities to misuse or misunderstand the concepts inherent in the manipulative.

As practical guidance for teachers, when extended practice with a manipulative is not possible due to practical or other limitations, teachers should consider introducing the abstract concepts by providing some support, *but not the most,* and build instruction from there to allow

students to reflect on understanding the idea rather than getting distracted or fixated on the manipulative itself. Likewise, instructional designers should consider that the amount of support being offered via the environment is not necessarily better and apparent engagement may not necessarily reflect that learning will take place. Moreover, as was observed in the current study, engagement at any level may give students (and perhaps teachers, too) the illusion of learning when in fact the learners might not be actually learning much.

Limitations and Future Directions

In spite of the strengths of this controlled experiment, there are some limitations that are important to note. Although this study employed random assignment, it did not explicitly examine students' prior experiences with the *balance analogy*. For example, it did not examine instances from playing on the see-saw or the using the term "balance" in their day-to-day lives. Future investigations should attempt to collect this information in case background knowledge or previous experience might impact receptivity to the manipulative or analogy. I provide a cautionary note along with this suggestion: asking students about their previous experiences with balancing may also sensitize students towards the instruction and therefore needs to be carefully handled. By accounting for previous experiences with analogies of balancing, one could explore whether students with this experience might respond more positively to interventions that include this rather than other interventions. Positive findings will add to existing theories on the role and effective use of analogies to understand mathematics (Lakoff & Johnson, 2003; Lakoff & Nunez, 2001).

A second limitation of the study was that some students were excluded from the analyses, leading to fewer third graders in the study than second graders. These students present an interesting case: why were these students perceiving the equal sign relationally despite going

through the same instruction in the classroom as their peers? Was it something they attended to in the classroom and understood, and, if so, might it be possible that other students had remnants of such instruction that gave them a head start in understanding the new ideas given in a multimedia instructional environment?

Another limitation was that this study did not examine what students were paying attention to during the intervention phase. Tracking students' eye movements in the four instructional conditions would offer information on which features were distracting and at which moments during the instruction, allowing researchers, designers and teachers to understand learning from these interventions in more depth.

Conclusions

In sum, this work provides evidence that students' understanding of the equal sign as an operational symbol can be disrupted even at an age when these patterns are considered to be entrenched by offering support through a brief intervention that directly ties the symbol with its meaning. This work also supports the idea that manipulatives with maximum support may not promote learning as well as manipulatives that are not as rich. From this, I argued that rich perceptual manipulatives, while intuitive, may constrain mathematical learning by not allowing the learner to extract and transfer the idea inherent in the manipulative to be used successfully after the manipulative is removed, when solving arithmetic problems. This study thus contributes to perspectives on the potential of cognitive distraction and hindrance caused by certain manipulatives, supporting theories of physically distributed learning and questioning those that consider the use of concrete manipulatives necessary for bridging the gap between concrete to abstract understandings in mathematics. Based on evidence, this work posits that rich support may not lead to optimal learning. Specifically, while the balance scale might foster learning,

forcing learners to make sense of the ideas of equivalence rather than showing them exactly how it works on a balance scale may be most conducive to learning, even when they are introduced in brief instruction, which is typically the case in many classrooms. In sum, providing opportunities for young students to wrestle with equivalence when it comes to balancing both sides of an equation—without doing all of the work for them—may be just the right balance for helping students learn.

REFERENCES

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and instruction*, *16*(3), 183-198.

Ball, D. L. (1996). Teacher learning and the mathematics reform: What we think we know and what we need to learn. *Phi Delta Kappan*, *77*(7), 500.

Baroody, A. J., & Coslick, R. T. (1998). *Fostering children's mathematical power: An investigative approach to K-8 mathematics instruction*. Mahwah, NJ: Erlbaum.

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

Baroody, A. J. (2006). Mastering the Basic Number Combinations. *Teaching children mathematics*, 23.

Behr, M., Erlwanger, S., & Nichols, E. (1976). *How children view equality sentences.* Tallahassee: Florida State University.

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

Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: Real problems, potential solutions. *Child Development Perspectives, 3*(3), 160-164.

Capraro, M. M., Ding, M., Matteson, S., Capraro, R. M., & Li, X. (2007). Representational implications for understanding equivalence. *School Science and Mathematics*, **107**, 86 – 88.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic & algebra in elementary school*. Portsmouth, NH: Heinemann.

Clark, R. C., & Mayer, R. E. (2011). E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. Pfeiffer.

Clements, D. H., & Battista, M. T. & Sarama, J. (2001). Logo and geometry. *Journal for Research in Mathematics Education Monograph Series.*

Clements, D. H., & Sarama, J. (2002). The role of technology in early childhood learning. *Teaching Children Mathematics*, *8*(6), 340.

Clements, D. H., & Sarama, J. (2004). Building Blocks for early childhood mathematics. Early *Childhood Research Quarterly, 19*, 181-189.

Dorward, J. (2002). Intuition and research: Are they compatible. *Teaching Children Mathematics*, *8*(6), 329-332.

Falkner, K., Levi, L., & Carpenter, T. (1999, December). Children's Understanding of Equality: A Foundation for Algebra. *Teaching Children Mathematics*, pp. 232-236.

Freiman, V., & Lee, L. (2004). Tracking primary students' understanding of the equality sign. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 415-422). Bergen, Norway: Bergen University College.

Fyfe, E. R., & McNeil, N. M. (2009). Benefits of "concreteness fading" for children with low knowledge of mathematical equivalence. *Poster presented at the Cognitive Development Society, San Antonio, TX*.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review,* 26, 9-25.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology, 96*, 424-436. doi:10.1037/0022-0663.96.3.424

Hattikudur, S., & Alibali, M. (2010). Learning about the equal sign: Does comparing with inequality symbols help? *Journal of Experimental Child Psychology*, *107(1)*, 15-30.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. Grouws (Ed.), Handbook for research on mathematics teaching and learning (pp. 65–97). New York: MacMillan.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2009) Concrete instantiations of mathematics: A double-edged sword. *Journal for Research in Mathematics Education, 40,* 90-93

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology*, *105*(2), 351.

Kaput, J. (1989). Linking representations in symbol system of algebra. In C. Kieran & S. Wagner (Eds.), *A Research Agenda for the Learning and Teaching of Algebra*. Hillsdale, NJ: Lawrence Erlbaum.

Kennedy, L.M. & Tipps, S. (1994). *Guiding children's learning of mathematics*, 7th ed. Belmont, CA: Wadsworth.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, *12(3),* 317-326.

Knuth, E., Stephens, A., McNeil, N., & Alibali, M. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37(*4), 297-312.

Lakoff, G., & Nunez, R. (2001). *Where mathematics come from: How the embodied mind brings mathematics into being*. Santa Fe Institute.

Lindgren, R., & Moshell, J. (2011) Supporting children's learning with body-based metaphors in a mixed reality environment. *Proceedings of the Interaction Design and Children Conference,* 177-180.

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

Mann, R. L. (2004). Balancing act: The truth behind the equals sign. *Teaching Children Mathematics*, 65-69.

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology, 104,* 1–21.

Mathews, P., Rittle-Johnson, B., McEldoon, K., & Taylor, R. (2012). Measure for measure: What combining diverse measures reveals about children's understanding of the equal sign as an indicator of mathematical equality. *Journal for Research in Mathematics Education, 43*(3), 316-350.

Martin, T. (2009). A theory of physically distributed learning: How external environments and internal states interact in mathematics learning. *Child Development Perspectives, 3*, 140–144.

Mayer, R. E. (2005c). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 183–200). New York: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd Ed). New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43-52.

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

McNeil, N. (2007). U-shaped development in math: 7-year-olds outperform 9-year olds on equivalence problems. *Developmental Psychology*, 43(3), 687-695.

McNeil, N. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79*(5), 1524-1537.

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

McNeil, N., & Alibali, M. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6(2)*, 285-306.

McNeil, N., & Alibali, M. (2005a). Knowledge change as a function of mathematics experiences: All contexts are not created equal. *Journal of Cognition and Development*, *6(2)*, 285-306.

McNeil, N. & Alibali, M. (2005b, July/August). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Developmental*, *76*(4), 883-899.

McNeil, N., Chesney, D., Matthews, P., Fyfe, E., Petersen, L., Dunwiddie, A., Wheeler, M. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of mathematical equivalence. *Journal of Educational Psychology, 104(4),* 1109-1121.

McNeil, N., & Fyfe, E. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction, 22,* 440-448.

McNeil, N., Fyfe E., Petersen L, Dunwiddie A., & Brletic-Shipley H. (2011, September/October). Benefits of practicing 4 = 2 + 2; Nontraditional problem formats facilitate children's understanding of mathematical equivalence. *Child Development*, *82*(5), 1620-1633.

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

McNeil, N. M., & Uttal, D. H. (2009). Rethinking the use of concrete materials in learning: Perspectives from development and education. *Child Development Perspectives, 3,* 137-139.

Molina M., & Ambrose, R. (2006). What is that equal sign doing in the middle?: Fostering relational thinking while negotiating the meaning of the equal sign. *Teaching Children Mathematics, 13*(2), 111-117.

Moyer-Packenham, P. S. (2010). Teaching mathematics with virtual manipulatives. Rowley, MA: Didax.

Moyer, P.S., Bolyard, J.J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching children Mathematics, 8(6),* 372-377.

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

National Council of Teachers of Mathematics (2014). *Illuminations: Resources for teaching math.* Retrieved Dec 2, 2014 from http://illuminations.nctm.org/.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2012). *Common Core State Standards for Mathematics*. Washington, DC: Authors.

National Library of Virtual Manipulatives (NLVM) Website. (2014). Retrieved Dec 2, 2014 from: http://nlvm.usu.edu/en/nav/vlibrary.html

National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*, U.S. Department of Education: Washington, DC..

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6,* 449-468.

Perry, M., Berch, D., & Singleton, J. L. (1995). Constructing shared understanding: The role of nonverbal input in learning contexts. *Journal of Contemporary Legal Issues, 6,* 213-236.

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

Petersen, L. A., & McNeil, N. M. (2013). Using perceptually rich objects to help children represent number: Established knowledge counts. *Child Development,* 84, 1020-1033.

Piaget J. (1970). *Science of education and the psychology of the child*. New York: Orion Press.

Powell S. (2012, June). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal*, 112(4), 627-648.

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

Rittle-Johnson, B., & Alibali, M. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, 91(1), 175-189.

Rittle-Johnson, B., Matthews, P. G., Taylor, R. & McEldoon, K. (2011). Assessing knowledge of mathematical equivalence: A construct modeling approach. *Journal of Educational Psychology*, 103, 85-104.

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

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

Stephens, A., Knuth, E., Blanton, M., Isler, I., Gardiner, A., & Marum, T. (2013). Equation structure and the meaning of the equal sign: The impact of task selection in eliciting elementary students' understandings. *Journal of Mathematical Behavior 32*(2), 173-182.

Van de Walle, J. (2012). *Elementary and Middle School Mathematics* (8th ed.). Boston, MA: Pearson Education Inc.

Uttal D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18,* 37-54. doi:10.1016/S0193-3973(97)90013-7

Uttal, D., O'Doherty, K., Newland, R., Hand, L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives, 3*, 156–159.

APPENDIX A

Table A1

Wilcoxon Test Results on Immediate Posttest and Delayed Posttest Scores by Instructional Condition

Problem Type	Instructional Condition							
	Balance Scale		Verbal Analogy		Core Idea		Control	
	V	*p*	V	*p*	V	*p*	V	*p*
Practice	89.0	0.82	92.5	0.07	89.0	0.82	27.0	0.35
Transfer	22.0	0.60	52.5	1.00	46.0	0.60	3.5	0.71
Practice + Transfer	73.0	0.60	193.5	0.87	114.5	0.99	22.0	0.34

APPENDIX B: SCREENING QUESTIONS

1. For each number sentence below, decide if the number sentence is True. In other words, does it make sense*?

After each problem, circle True, False, or Don't Know.

Examples:

3 + 4 = 7 (True) False Don't Know

3 + 4 = 12 True (False) Don't Know

---

a) 8 = 8 True False Don't Know

b) 7 + 6 = 0 True False Don't Know

c) 5 + 3 = 3 + 5 True False Don't Know

d) 8 = 5 + 10 True False Don't Know

e) 3 + 1 = 1 + 1 + 2 True False Don't Know

f) 4 = 4 + 0 True False Don't Know

2. Here are some addition problems. Solve them as best as you can. Write your answer in the blank.

a) 4 + 5 = _____

b) 13 + 15 = _____

c) 24 + 19 = _____

d) 3 + 3 + 3 = _____

* These questions are from Rittle-Johnson et al, 2011.

3. Here are some subtraction problems. Solve them as best as you can. Write your answer in the blank.

a) 6 – 4 = _____

b) 18 – 13 = _____

c) 26 – 18 = _____

d) 9 – 3 – 3 = _____

APPENDIX C: IDENTIFYING THE TWO SIDES OF THE EQUATION

Pretest

1. This problem has two sides. Circle the choice that correctly breaks the problem into **two sides**[*]:

$$5 = 3 + 2$$

a)

Side A	Side B
5 =	3

b)

Side A	Side B
5	3 + 2

c)

Side A	Side B
5 =	3 + 2

d)

Side A	Side B
5 = 3 + 2	3 + 2 = 5

e) ?

2. This problem has two sides. Circle the choice that correctly breaks the problem into **two sides**[*]:

$$4 + 1 = 2 + 3$$

a)

Side A	Side B
4 + 1 =	2

b)

Side A	Side B
4 + 1= 2	+ 3

c)

Side A	Side B
4 + 1 = 2	4 + 1 = 2 + 3

d)

Side A	Side B
4 + 1	2 + 3

e) ?

[*] These questions are adapted from the questions by Rittle-Johnson et al., 2011.

Posttest Questions

1. This problem has two sides. Circle the choice that correctly breaks the problem into **two sides**[*]:

$$2 + 3 + 5 = 9 + 1$$

a)

Side A	Side B
2 + 3 +	5 = 9 + 1

b)

Side A	Side B
2 + 3 + 5 + 9	= 1

c)

Side A	Side B
2 + 3 + 5	9 + 1

d)

Side A	Side B
2 + 3 + 5 = 9 + 1	1 + 9 = 5 + 3 + 2

e) ?

[*] These questions are adapted from the questions by Rittle-Johnson et al., 2011.

2. This problem has two sides. Circle the choice that correctly breaks the problem into **two sides**[*]:

$$5 + 1 = 4 + 2$$

a)

Side A	Side B
5 + 1 =	4

b)

Side A	Side B
5 + 1 = 4	+ 2

c)

Side A	Side B
5 + 1 = 4 + 2	4 + 2 = 5 + 1

d)

Side A	Side B
5 + 1	4 + 2

e) ?

[*] These questions are adapted from the questions by Rittle-Johnson et al., 2011.

APPENDIX D: AFFECTIVE SURVEY

1. Do you think your friend would like playing this game? Yes No

   Why Yes? / Why No?

2. How often do you play on a computer or a laptop at home in a week?

   Never Special days Weekends Almost Everyday

3. Questions about the intervention:

Question				
I would like to play the game again	1	2	3	4
Math was fun in this game.	1	2	3	4
I think I was really good at playing this game	1	2	3	4
I liked solving math problems in the game	1	2	3	4
I would like to get more difficult math problems in the Marty game	1	2	3	4
I think playing this game helped me solve math problems	1	2	3	4

Note. 1=Absolutely Disagree, 2=Disagree, 3=Agree, 4=Absolutely Agree

4. Do you think it was a game or was it just math?