Individual Differences in the Relational Understanding of Mathematical Equivalence

Emmanuelle Adrien

A Thesis

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Child Study) at

Concordia University

Montreal, Quebec, Canada

September 2012

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Emmanuelle Adrien

Entitled: Individual Differences in the Relational Understanding of Mathematical Equivalence

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Child Study)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________________ Chair

Dr. Diane Pesco

______________________________________ Examiner

Dr. Jeffrey Bisanz

______________________________________ Examiner

Dr. Sandra Martin-Chang

______________________________________ Supervisor

Dr. Helena Osana

Approved by ____

Dr. Richard Schmid, Chair of Department

___________ 2012 ______________________________________

Dr. Brian Lewis, Dean of Faculty

ABSTRACT

Individual Differences in the Relational Understanding
of Mathematical Equivalence

Emmanuelle Adrien

Many children have a limited view of the equal sign and do not understand that the left and right sides of an equation need to represent the same amount, even after instruction. In the present study, I investigated domain-general and domain-specific factors that account for differences in children's performance on noncanonical equivalence problems. Fifty-six third- and fourth-grade students were asked to solve a series of noncanonical equivalence problems before and after receiving explicit instruction on the equal sign. As a group, students showed significantly higher scores after instruction, but close to a third of them were unsuccessful on more than half of the items on the posttest. Individual interviews were conducted to determine the source of the variability; and interview tasks assessed conceptual understanding of the equal sign in a nonsymbolic context and in a symbolic context (as measured by students' ability to rate and generate equal sign definitions and to justify their answers to noncanonical equations). Regression analyses indicated that of all the variables, only general ability and mathematical fluency were significant predictors of performance on the posttest. Qualitative analyses of the interview data revealed that students generated three different types of definitions of the equal sign (i.e., operational, relational, combined), and that there was a significant effect of definition type on posttest performance. No significant difference was found in the performance of students who offered combined and relational definitions, but as a group, they outperformed students who held entirely operational definitions. The study's limitations, implications of the research findings, and avenues for future research are discussed.

ACKNOWLEDGEMENTS

I would first like to express my gratitude to my supervisor, Dr. Helena Osana, without whom this thesis would not have become a reality. I truly admire your dedication to each one of your students and your desire to help them in their research endeavors. I greatly appreciated the many discussions we had while working on this thesis, which always left me energized and excited about my research. Thank you for your encouraging nature, patience, and flexibility. I am honored to be working with you, and I look forward to many more years of collaboration.

I owe sincere thankfulness to the members of my committee, Dr. Sandra Martin-Chang and Dr. Jeffrey Bisanz, for their insightful advice and continuous support. I also want to thank Rebecca Watchorn, Vanessa Rayner, and Jody Sherman-LeVos for the input on study design and measures they provided while working on the larger project of this study. It is a great pleasure to thank Allyson Cooperman, Katarzyna Przednowek, Stephanie Peccia and Amanda Corber who helped me tremendously during the data collection phase of the research. Thank you for your diligence and enthusiasm, and thank you for your moral support.

Words cannot express how much I am indebted to my husband Marc, my parents Jessica and Alix, my sister Noémie, my friends and the rest of my family. I am blessed to be surrounded by so many loving and caring people. Thank you for believing in me as much as you do and for always being there for me. This would have not been possible without all of you. I love you all very much.

Table of Contents

List of Figures

List of Tables

Table

Chapter 1: Statement of the Problem

When children experience difficulty in mathematics in elementary school, they are likely to fall behind and have a hard time meeting high school mathematics standards (Hammer, 2010). Students typically start algebra courses in high school, and academic success is often equated with success in algebra (Kaput, 2008). Many researchers now believe that algebraic thinking should be introduced in elementary school and that it could help students with the transition to the formal algebra curriculum (e.g., Carpenter, Franke, & Levi, 2003; Carraher, Schilemann, Brizuela, & Earnest, 2006; Kaput, 2008). One way to do so is to teach mathematical equivalence to young students.

Mathematical equivalence refers to a relation between two sides of an equation; the left side and the right side of the equal sign represent the same amount. Numerous researchers have reported that children perform poorly on nonstandard equivalence problems – that is, those that deviate from the standard form (i.e., a = b + c; Carpenter et al., 2003). Although some researchers have proposed different explanations for this finding, it is still unclear why such difficulties occur. At first glance, the equal sign might not seem like an important symbol to study, but it appears that many children do not understand what the equal sign really means (e.g., Carpenter et al., 2003; McNeil et al., 2006). This is an area worthy of exploration, as it has been reported that understanding the meaning of the equal sign is at the basis of algebraic thinking (Franke, Carpenter, & Battey, 2008). Algebra is often seen as the gateway to higher mathematics; therefore, if children fail to grasp algebraic concepts, it could affect their future academic achievement and eventual choice of career (Stephens, 2006; Uttal, O'Doherty, Newland, Hand, & DeLoache, 2009).

The present study examined children's individual characteristics that could explain their struggle with equivalence problems. Several studies have examined external factors related to equivalence. For example, a few researchers have proposed that the way children typically encounter the equal sign in their textbooks and during instruction – with the operators on the left side of the equal sign and "the answer" on the right side – might make it difficult for them to view the equal sign as a relational symbol (e.g., McNeil et al., 2006; Seo & Ginsburg, 2003). Researchers interested in finding the most effective way to teach mathematical equivalence have yet to agree on one specific method for doing so. Some contend that instruction through classroom discussions might be effective (Carpenter et al., 2003; Sáenz-Ludlow & Walgamuth, 1998), while others have examined the effects of explicit and direct explanations of concepts and procedures on performance on equivalence problems (Watchorn, Osana, Sherman, Taha, & Bisanz, 2011).

Other studies have described the cognitive factors that influence children's achievement in different mathematical areas. Working memory, nonverbal intelligence, and conceptual understanding of specific mathematical concepts have all been linked to mathematics achievement (e.g., Alloway & Passolunghi, 2011; Canobi, 2004; Fuchs et al., 2010). More research is needed to identify the cognitive factors that specifically affect children's understanding of equivalence, however. Gaining a better understanding of the cognitive factors that influence children's relational understanding of mathematical equivalence will be valuable in many ways. Researchers and program developers might use this information to create more effective instructional methods for teaching equivalence. Teachers will become more knowledgeable of their students' specific needs, and this will allow them to adapt their teaching to accommodate these needs.

Chapter 2: Literature Review

Mathematical Equivalence

At a young age, children are introduced to different mathematical symbols and are expected to learn their meaning. The equivalence relationship represented by the equal sign is one of the basic concepts in formal arithmetic (Baroody & Ginsburg, 1983), and it is used in mathematical equations to signify that the quantity represented on its left side is the same as the quantity represented on its right side. It is important for young children to hold an accurate understanding of the meaning of the equal sign if they are to succeed in more advanced mathematics when they get older (Hammer, 2010; McNeil & Alibali, 2005a). To illustrate, one procedure that is often taught in algebra is to "cancel out" like items on both sides of the equation. Children who understand the equal sign will understand the meaning of this procedure. Consider the equation 5 + ___ + 169 = 3 + 8 + 169. A child with an understanding of the relationship between the amounts on both sides of this equation would know to ignore the 169 on both sides of the equation (i.e., "cancel them out") and concentrate on solving the simpler problem (i.e., 5 + ___ = 3 + 8).

Falkner, Levi, and Carpenter (1999) described the experience of a first- and second-grade teacher who realized that her students did not understand what the equal sign meant; this was evident to her because none of her students were able to solve equivalence problems such as 8 + 4 = __ + 5 correctly. After extensive instruction on the meaning of the equal sign, she found that her students began to solve equivalence problems accurately, and that they were also able to understand more sophisticated algebraic problems (e.g., understanding that a is larger than b in the sentence $a = b + 2$).

Knuth, Stephens, McNeil, and Alibali (2006) investigated the relationship between middle school students' understanding of the equal sign and their performance on algebraic equations. They gave their participants a paper-and-pencil test on which one item required them to explain what the equal sign meant and another item asked them to solve two algebraic equations (e.g., $3m + 7 = 25$). Knuth et al. found that the majority of the middle school students participating in the study did not view the equal sign as a symbol of mathematical equivalence. Furthermore, they found that the students who understood the meaning of the equal sign were more likely to solve the equations correctly than the students who did not understand the meaning of the equal sign.

In the mathematical equivalence literature, authors commonly refer to two types of equations: canonical and noncanonical (e.g., Gilmore & Bryant, 2006; Sherman & Bisanz, 2009). In canonical equations (e.g., a + b = c), there are two or more numbers accompanied by operators on the left side of the equal sign and only one number on the right side of the equal sign. An equation without any operator (e.g., 3 = 3) would be an example of a noncanonical equation. Other examples of noncanonical equations would include all equations where there are more than one number and at least one operator on the right side of the equal sign. In their study, Sherman and Bisanz used four types of symbolic representations of the equal sign in noncanonical contexts: (a) identity problems (a + b = a + __), (b) commutativity problems (a + b = b + __), (c) part-whole problems (a + b = c + __), and (d) combination problems (a + b + c = a + __).

Children's Thinking About Equivalence

Even though the equal sign is central to mathematics instruction, many elementary school children tend to interpret it as a signal to "do" something, rather than as a symbol used to show equivalence (Kieran, 1981). Children who hold the former interpretation of the equal sign are said to exhibit an "operator view," and children who hold the latter interpretation of the equal sign are said to exhibit a "relational view" (e.g., Baroody & Ginsburg, 1983; McNeil & Alibali, 2005a; Sáenz-Ludlow & Walgamuth, 1998; Seo & Ginsburg, 2003). When children see the equal sign as an operational symbol, they are exhibiting a less sophisticated form of thinking compared to those who see the equal sign as indicating an equivalent relation between two quantities (McNeil & Alibali, 2005a). McNeil and Alibali (2005b) described three "operational patterns" that can be observed in children: (a) the idea that the equal sign and the unknown come together at the end of a problem, (b) the idea that the equal sign signifies calculating the total, and (c) the idea that mathematics problems are solved by performing all given operations on all given numbers. When children have an operator view of the equal sign, they tend to reject noncanonical equations and thus have difficulty correctly determining the unknown in equations such as 8 + 4 = __ + 5 (Seo & Ginsburg, 2003).

Several studies (e.g., McNeil & Alibali, 2004; Watchorn & Bisanz, 2005) have found that school-aged children have a low success rate on equivalence problems if the problems are not in canonical form. The operational manner in which children view the equal sign leads them to solve equivalence problems in ways that reflect specific misconceptions. Carpenter et al. (2003) explained that when given problems similar to 8 + 4 = __ + 5, children typically respond in one of five ways. First, some children might display an "answer comes next" conception and would answer that the missing number in this case should be 12. When these children see the equal sign, they expect the answer to the operation on its left side to appear on its right side and they will disregard any operations that appear on the right side. This conception relates to McNeil and Alibali's (2005b) first operational pattern described above. Second, some children might exhibit a "use all the numbers" conception and would answer that the missing number should be 17. These children believe they have to use all the numbers and all the operations to solve the given equation. This conception was referred to in the above description of McNeil and Alibali's third operational pattern. Third, some children might show an "extend the problem" conception and would write 12 on the line, but add a second equal sign after the 5 and write the number 17 to the right of it (i.e., 8 + 4 = _12_ + 5 = *17*). These children, like those who display an "answer comes next" conception, believe the answer to the operation of the left side of the equal sign needs to appear on its right side, but they also want to take into account the "+ 5".

There are two ways children who have a relational view of the equal sign might solve the equation written above. They might calculate the sum on the left side of the equal sign (i.e., 12) and compute the missing number on the right side of the equal sign by performing the necessary operation (in this case, subtraction: 12 – 5 = 7). Other children might use a more abstract way to solve the problem and recognize the relationship between the numbers on both sides of the equal sign. Since 5 (on the left side) is one more than 4 (on the right side), the missing number needs to be one less than 8 (therefore, 7).

Factors Affecting Children's Thinking about the Equal Sign

Baroody and Ginsburg (1983) reported that earlier studies on children's understanding of the equal sign reflected one of two views. The first view contends that children's tendency to view the equal sign in an operational manner is due to their prior knowledge in arithmetic. The second view suggests that children's flawed understanding of the equal sign results from developmental limitations. Baroody and Ginsburg hypothesized that if the first view were true, students receiving instruction that promoted viewing the equal sign as a relational symbol would view it as such. If the second view were true, then the type of instruction the students received would not matter because they lacked the cognitive abilities necessary for them to understand the relational meaning of the equal sign.

Baroody and Ginsburg (1983) recruited students from grades 1, 2, and 3. These students attended a school where the mathematics curriculum emphasized a relational view of the equal sign. The authors asked participants to evaluate if given equations were correct ("made sense") or not. The equations were of canonical and noncanonical types. Although they found that students' judgments were inconsistent, they also found that more than half accepted as correct noncanonical equations of a type to which they had never been exposed. Thus, Baroody and Ginsburg rejected the view that children cannot understand the meaning of the equal sign because of developmental limitations. More recent research has confirmed this finding; several researchers have shown that young children have the cognitive ability to understand the equal sign relationally. It is widely accepted that children are able to reason in a relational manner before algebra is formally taught to them (Sherman & Bisanz, 2009), even as early as grade 1 (Carpenter et al., 2003; Carpenter & Levi, 2000).

If young children are cognitively able to understand the meaning of the equal sign, why do they typically solve equivalence problems using an incorrect operational view? It is possible that external factors, such as classroom instruction (including specific tasks and teacher practice) and the ways in which the equal sign is presented in curricular materials (e.g., textbooks and teacher materials) may explain the phenomenon. It is also possible that individual differences, such as prior knowledge and working memory, are at play. In the following section, I review some of the literature that examines possible external influences and individual cognitive factors involved in children's understanding of the equal sign.

External factors that influence children's thinking about the equal sign. Sherman and Bisanz (2009) hypothesized that the type of problem may differentially affect children's performance on equivalence problems. They predicted that identity (a + b = a + __) and commutativity (a + b = b + __) problems might be easier to solve than part-whole (a + b = c + __) or combination (a + b + c = a + __) problems because they can be solved without computation; children simply have to notice that both sides of the equal sign have the same addends. The study's results support this hypothesis: the authors explained that children's performance on equivalence problems followed one of three patterns. One child (out of their sample of 24 children) appeared to understand equivalence and solved a variety of problems accurately. Some children (25%) demonstrated a less sophisticated understanding and were only able to correctly solve identity or commutativity types of problems. The other children (71%) were not able to solve any equivalence problem accurately. Thus, it appears that students' performance might also be dependent on the type of task they are asked to complete.

The types of equations found in mathematics textbooks might be another reason children tend to hold an operational view of the equal sign (Seo & Ginsburg, 2003). McNeil et al. (2006) examined the content of several middle-school (grades 6-8) mathematics textbook series to determine the various contexts in which the equal sign was presented. They identified one "standard" context where equations had operations on the left side of the equal sign and "the answer" on the right side (e.g., 3 + 4 = 7). They also identified four "nonstandard" contexts: (a) equations with operations on both sides of the equal sign (e.g., 5 + 2 = 3 + 4), (b) equations with operations only on the right side of the equal sign (e.g., 7 = 3 + 4), (c) equations without explicit operations on either side (e.g., 7 = 7), and (d) no equation (e.g., using <, >, or = to complete a statement such as 5 __ 8). The authors found that very few of the middle-school textbooks they reviewed presented equations that had operations on both sides of the equal sign.

Following the textbook analyses, McNeil et al. (2006) conducted experiments with middle-school students in which they assigned each participant to view the equal sign in a specific way (i.e., reflexive, operations equal answer, operations on the right side, operations on both sides). They found that students were more likely to exhibit a relational understanding of the equal sign when it was presented in a nonstandard equation context than when it was presented in the context of canonical equations. Repeatedly seeing the equal sign presented in canonical equations might reinforce children's operational view. It is therefore problematic that textbooks rarely present equations with operations on both sides of the equal sign.

Li, Ding, Capraro, and Capraro (2008) compared sixth-grade Chinese students' interpretations of the equal sign to those of sixth-grade American students. The participants were asked to solve three equivalence problems (e.g., 6 + 9 = __ + 4) and to judge whether a noncanonical equation (e.g., 6 + 8 = 3 + 11) was true or false. Li et al. found that the Chinese students clearly outperformed the American students (98% correct vs. 28% correct across all four items). The authors also compared the use of the equal sign in Chinese and American sixth-grade mathematics textbooks and found that similarly to what McNeil et al. (2006) had reported, the equal sign was presented more often using different types of equations and in different contexts in Chinese textbooks compared to American textbooks.

Li et al. (2008) also examined textbooks used in mathematics methods classes for preservice teachers in the United States and teachers' guidebooks in China. They found that only two of the six American textbooks reviewed directly addressed the equal sign. Furthermore, none of the textbooks provided examples of lessons or indications on how the equal sign should be taught to children. In contrast to the American materials, the Chinese teaching materials demonstrated lessons aimed at developing children's understanding of the equal sign as a relational symbol, and teachers were encouraged to use the equal sign in problems of various types and in different contexts with their students.

The findings that Chinese students performed better than American students on equivalence problems combined with the differences in instructional materials between the two countries suggest that instruction also plays a role in children's conceptions about the equal sign. Knowing that, it is worrisome to think that teachers in North America spend little or no time on explicit instruction about the meaning of the equal sign (McNeil & Alibali, 2005a; McNeil et al., 2006).

Another body of literature points to the types of instruction that may enhance children's view of the equal sign. In the first place, several studies (e.g., Seo & Ginsburg, 2003) have reported that the meaning of the equal sign is typically not taught explicitly in school. A number of researchers have therefore investigated the effectiveness of different methods of teaching children to think relationally. Carpenter et al. (2003) suggested an approach to teaching children equivalence based on children's thinking and classroom dialogue. In this instructional context, the teacher presents students with mathematical problems that might elicit children's misconceptions and encourage them to verbalize their views about the equal sign. The teacher would challenge the students' misconceptions with the goal of helping them develop a relational understanding of the equal sign.

In a study using an experimental design, Watchorn et al. (2011) provided grade 2 and grade 4 students with one of four types of instruction on how to solve equivalence problems (i.e., conceptual without manipulatives, conceptual with manipulatives, procedural without manipulatives, procedural with manipulatives). The authors were interested in determining the influence of the different types of instruction on students' performance on equivalence problems and on students' understanding of equivalence. No instructional condition emerged as clearly superior relative to the others, but the authors nevertheless found that some students were still unable to correctly solve equivalence problems, even immediately following instruction. Thus, despite attempts to teach relational thinking, there is still variability in children's views of the equal sign, even after instruction. Therefore, I hypothesize that such individual differences may be explained by specific cognitive factors. A number of studies and reviews of literature point to possible predictors, which are addressed in the following section.

Cognitive predictors of mathematical performance. The studies reviewed above, all relate to external factors influencing children's understanding of the equal sign.

It is equally important to consider the individual characteristics that can affect children's understanding of mathematical concepts, specifically of the equal sign. Considering that few studies have explored this issue, I will review literature related to broader mathematical concepts.

Fuchs et al. (2010) reported that, in contrast to the numerous studies on domain-specific and domain-general competencies related to reading skills, little is known about the general or specific competencies that influence students' learning of mathematics in school. The authors mentioned that studies to date on the relationship between children's basic numerical cognition and domain-general abilities do not provide strong conclusions about the importance of each domain of competence for mathematics learning. They examined the contributions of domain-general abilities (i.e., language, nonverbal problem solving, working memory, attentive behavior, and processing speed) and basic numerical cognition (i.e., linking quantities to the appropriate Arabic numeral, placing numbers on a number line marked only at 0 and 100) of first-graders on their mathematics development. The mathematics outcomes Fuchs et al. were interested in were achievement on procedural calculations (i.e., double-digit addition and subtraction problems with and without regrouping) and on word problems. They found that basic numerical cognition explained 26.5% of the variance in growth when it came to the procedural calculation problems. The domain-general abilities did not make a significant contribution.

For the word problems, Fuchs et al. (2010) found that basic numerical cognition and domain-general abilities together explained 32% of the variance in growth and that they did so in comparable proportion. The authors identified four domain-general abilities as uniquely predictive of word problem solving: working memory, nonverbal problem solving, language, and teachers' ratings of attentive behavior. While this study did not specifically measure performance on equivalence problems, one may hypothesize that domain-general factors (such as working memory) and domain-specific factors (such as basic number knowledge and conceptual understanding of the equal sign) could explain the variance in performance on equivalence problems.

In Baddeley and Hitch's (1974; as cited by Rasmussen & Bisanz, 2005) model, working memory is defined as a system that involves short-term storage and manipulation of information that is needed to perform different cognitive tasks. There are three main components to this model. First, there is the central executive, which plays a role in the regulation, control, and monitoring of cognitive processes; there are also two subsystems that are responsible for temporarily storing different kinds of information: the phonological loop, which handles auditory and verbal information, and the visuo-spatial sketchpad, which handles visual and spatial information (De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière, 2009; Meyer, Salimpoor, Wu, Geary, & Menon, 2010; Raghubar, Barnes, & Hecht, 2010).

The literature describes the role of working memory in mathematical learning in two ways. First, working memory (especially the central executive and the phonological loop components) is involved in the acquisition of new knowledge and complex skills, and also in the acquisition of new solutions strategies (Gathercole, Pickering, Knight, & Stegman, 2003; Holmes & Adams, 2006). A child with a large working memory may find it easier to acquire new knowledge and procedures than a child who has a more limited working memory (LeFevre, DeStefano, Coleman, & Shanahan, 2005). Second, working memory is involved during the performance of a mathematical task in terms of the processing and storage of information (Gathercole et al., 2003). More specifically, the visuo-sketchpad has been associated with the encoding, retention, and manipulation of numbers during mental calculations, and the central executive component of working memory is believed to be important for choosing between previously learned strategies during problem solving (Holmes & Adams, 2006).

Alloway and Passolunghi (2011) examined the contribution of working memory and verbal ability to the mathematical skills of seven- and eight-year-old Italian children. They used the Automated Working Memory Assessment (AWMA) as a measure of working memory. The AWMA assesses verbal and visuo-spatial short-term and working memory. As a measure of mathematical ability, they administered a test that had four components: (a) number operations, (b) quantity discrimination, (c) number production, and (d) number ranking. They used a second measure of mathematical ability that specifically assessed computational skills. Alloway and Passolunghi found that working memory was indeed related to mathematical skills, but that the relationships differed for each age group depending on the types of memory task and mathematical skill. For instance, for the seven-year-olds, verbal memory predicted performance on the number ranking task and on the arithmetic skills measure. For the eight-year-olds, it was visuo-spatial memory that predicted these same skills.

In their longitudinal study, Krajewski and Schneider (2009) investigated the importance of domain-specific mathematical precursors (i.e., basic numerical skills and quantity-number concepts) and of non-specific cognitive precursors (i.e., nonverbal intelligence, phonological memory capacity, speed of access to long-term memory, and socioeconomic status) to mathematical achievement in school. The authors measured these precursors when their German participants were at the end of kindergarten. They assessed children's mathematical achievement when they were at the end of first grade and again at the end of fourth grade using standardized German mathematics tests. Krajewski and Schneider developed models to explain the interplay of their various predictors and mathematical achievement. They found that the domain-specific mathematical competencies assessed in kindergarten predicted later mathematical achievement better than non-specific cognitive factors. They also found, however, that nonverbal intelligence predicted the specific mathematical precursors and thus indirectly affected mathematical achievement in school. Speed of access to long-term memory and socioeconomic status were also related to mathematical achievement.

In their study, Jordan, Glutting, and Ramineni (2010) were interested in examining the unique contribution of children's symbolic number sense to mathematics achievement. They gave first grade students a number sense measure (i.e., a domain-specific measure) that assessed their counting knowledge, number recognition, ability to compare and combine numbers, nonverbal calculation skills, and capability to solve story problems. They also gave these students general cognitive measures (vocabulary, spatial reasoning, and short-term and working memory) and used them as control variables in their analyses. Jordan et al. assessed the participants' mathematics achievement with the Woodcock-Johnson III (WJ-III) towards the end of first grade and again towards the end of third grade. They found that the domain-specific measure (i.e., number sense) was significantly related to all other variables. This led the authors to conclude that number sense is a strong predictor of later mathematics outcomes.

Successfully solving equivalence problems involves two types of knowledge (e.g., Rittle-Johnson & Alibali, 1999; Rittle-Johnson & Koedinger, 2009). Conceptual knowledge is required to understand that the equal sign is a relational symbol, and procedural knowledge is needed to perform any specific calculations (e.g., addition or subtraction) necessary to solve the problem. Several studies have examined children's conceptual understanding of different mathematical concepts, and more specifically, Canobi (2004) explored individual differences in children's addition and subtraction knowledge. The results of her study suggest that there are individual differences in students' procedural (e.g., using different procedures flexibly to solve problems) and conceptual knowledge (e.g., showing an understanding of commutativity and part-whole relations). She also found that conceptual ability was not related to grade level, and explained that this might be due to an educational emphasis placed on solving problems successfully rather than on understanding the underlying concepts.

Gilmore and Bryant (2006) investigated the relationship between children's conceptual understanding of the principle of inversion (i.e., the inverse relationship between addition and subtraction) and their arithmetical skills as defined by procedural computation used in solving canonical and noncanonical equations. They presented children with canonical and noncanonical equations and found that for a majority of children, conceptual understanding was related to their computation skills. There were some children for whom conceptual understanding was greater than their computation skills, however. The authors believed that these results supported the idea that considering individual differences in children's conceptual understanding of mathematical concepts is important for mathematics instruction.

In the context of equivalence, Rittle-Johnson and Alibali (1999) investigated the relationship between conceptual knowledge of equivalence and procedural knowledge in fourth- and fifth-graders. Specifically, they wanted to see the impact conceptual instruction would have on children's procedures for solving equivalence problems and the impact procedural instruction would have on their conceptual understanding. They found that conceptual and procedural knowledge influenced each other. More specifically, the authors found that the conceptual understanding of children who received conceptual instruction increased, but so did their procedural knowledge. Children who received procedural instruction saw their procedural knowledge improve and also showed an increased conceptual understanding.

Children can hold relational and operational views of the equal sign at the same time, but will adhere to one or the other depending on the context in which the equal sign is presented (Seo & Ginsburg, 2003). In their study, Seo and Ginsburg found that children's views varied depending on the context in which the equal sign was presented. Specifically, students interpreted the equal sign in an operational manner when the equal sign was presented symbolically, in both canonical and noncanonical equations, but they gave a relational interpretation when the equal sign was presented with Cuisenaire rods or in a money context.

These findings led some researchers to explore children's views of the equal sign in specific contexts. In their study, Sherman and Bisanz (2009) wanted to know if children would have different success rates if they solved equivalence problems either symbolically (e.g., 5 + 2 = 4 + __) or nonsymbolically (i.e., using manipulatives). The second-grade children who participated in their study were more successful when solving nonsymbolic equivalence problems than equivalence problems presented symbolically. Those who were in the nonsymbolic group were also more likely to justify their answers using relational reasoning than those in the symbolic group. The authors concluded that children may encode equivalence problems in distinct ways (with or without symbols). Thus, it appears that children may not demonstrate proficiency on equivalence problems presented symbolically, but still possess a conceptual understanding of the equal sign.

Children's relational view of the equal sign in a nonsymbolic context will not necessarily transfer automatically to a symbolic context (Seo & Ginsburg, 2003), however. Uttal et al. (2009) similarly argued that, "the manipulation of concrete objects is not, in itself, enough to give children the ability to understand abstract, symbolic representations of mathematical ideas" (p. 156). Taken together, the findings described above suggest that there are individual differences in children's conceptual understanding of the equal sign in both nonsymbolic and symbolic contexts.

As a whole, the research on individual differences reviewed above suggests that domain-general abilities, such as working memory, relate to children's mathematics achievement in general (Alloway & Passolunghi, 2011; Fuchs et al., 2010; Krajewski & Schneider, 2009). There is also evidence to support the claim that domain-specific abilities (e.g., number sense) correlates with mathematics achievement (Jordan et al., 2010). Several authors believe that there are individual differences in terms of students' conceptual and procedural understanding in mathematics (Canobi, 2004; Gilmore & Bryant, 2006) and that the two types of knowledge might be related (Rittle-Johnson & Alibali, 1999). In the context of equivalence, results from two studies suggest that children's conceptual understanding of the equal sign might be apparent in some contexts and not in others, and in particular, they seem more likely to demonstrate a relational view of the equal sign in nonsymbolic contexts (Seo & Ginsburg, 2003; Sherman & Bisanz, 2009).

Present Study

The findings from several studies (e.g., Baroody & Ginsburg, 1983; Carpenter & Levi, 2000; Carraher et al., 2006) indicate that young children are able to develop a relational understanding of the equal sign. Not all children will achieve that goal with the same ease or at the same rate, however. In the present study, I investigated the student variables that may influence their relational understanding of mathematical equivalence.

This study was part of a larger study in which it was found that students' understanding of equivalence varied according to the type of instruction they had received (Watchorn et al., 2011). During the 2010-2011 school year and as part of the larger study, seven teachers took part in professional development (PD) workshops on the topic of mathematical equivalence, after which they implemented explicit instruction on the equal sign in their respective classrooms. Two participants were third-grade teachers and two others were fourth-grade teachers, and their students formed the sample for the present study. The purpose was not to test whether the students' performance on equivalence problems improved as a result of their teachers participating in PD workshops, but rather to understand what individual factors could explain any variance in students' performance after instruction.

For the present study, students' performance on equivalence problems was assessed before they received any kind of instruction on the subject. Students' performance on equivalence problems was assessed again after they had received formal instruction on the topic from their teachers. The students were given additional measures in class and during individual meetings to determine which factors could explain the trajectory of their performance from pretest to posttest.

I focused on children's cognitive factors rather than on the external factors that could explain individual differences in performance on equivalence problems because in the context of the larger study, these external factors were not controlled or manipulated. In order to predict mathematical outcomes, both domain-general and domain-specific factors are important to consider (Raghubar et al., 2010). In the present study, I explored the possible links between students' performance on equivalence tests and their domain-general abilities (i.e., working memory) and domain-specific abilities (i.e., conceptual understanding of the meaning of the equal sign in a symbolic and in a nonsymbolic context). Nonverbal intelligence was used as a control for general ability, and mathematical fluency was used as a measure of students' automaticity of addition and subtraction facts. The specific research questions were: (a) Will the domain-general factor (i.e., working memory) be a significant predictor of students' performance on equivalence problems? (b) Will domain-specific factors, their conceptual understanding of the equal sign in a nonsymbolic context and their conceptual understanding of the equal sign in a symbolic context (as measured by their ability to rate and generate equal sign definitions and to justify their answers to noncanonical equations), be significant predictors of students' performance on equivalence problems?

The literature reviewed has revealed links between domain-general abilities (such as working memory) and specific mathematics outcomes (such as performance on word problems; Fuchs et al., 2010). Working memory has been associated with mathematics learning and performance (Gathercole et al., 2003; Holmes & Adams, 2006). Different aspects of working memory (e.g., visuo-spatial and verbal), for example, appear to predict mathematics achievement (Alloway & Passolunghi, 2011). I therefore hypothesized that working memory abilities would predict performance on equivalence problems.

In addition, several studies have linked conceptual understanding to arithmetic performance (e.g., Gilmore & Bryant, 2006), and this led me to predict that performance on measures assessing students' understanding of the equal sign in a symbolic context, specifically the Rating Definitions and Generating Definitions tasks, would predict performance on equivalence problems. Since children seem more likely to express a relational understanding of the equal sign in nonsymbolic contexts regardless of their symbolic understanding (Seo & Ginsburg, 2003; Sherman & Bisanz, 2009), I hypothesized that, in contrast, performance on the Nonsymbolic Task would not be a significant predictor.

Examining the factors that predict children's understanding of the equal sign will make various theoretical and practical contributions. First, because "it is possible that a greater number of students actually understand the equal sign in a relational way, but they may be unable to demonstrate that understanding in an equation-solving situation because the equations they typically encounter in school frequently elicit the operational interpretation" (McNeil et al., 2006, p. 381), the results of the present study might contribute to the existing literature by shedding light on the relationship between conceptual knowledge and procedural knowledge in the context of equivalence. In addition, understanding the factors that influence students' understanding of equivalence might help teachers identify the potential causes of their students' mathematical difficulties, adapt their teaching strategies to better address students' specific needs, and inform them on the areas of their students' learning on which to focus.

Chapter 3: Method

Context

This study was part of a larger project in which seven elementary school teachers took part in three PD workshops during the 2010-2011 school year to learn about children's thinking regarding mathematical equivalence and to help create classroom activities on that topic. The ultimate goal of the project was to test these classroom activities on a large scale across Canada during a subsequent phase. The students in four of the teachers' classrooms formed the sample for the present study.

The teachers' first PD workshop took place on February 3, 2011, and it aimed to inform teachers on children's thinking about equivalence. During their first workshop, teachers learned about students' conceptions and misconceptions of equivalence. After the first workshop, the teachers engaged in some informal probing activities at least once in each of their classrooms to evaluate their own students' views on equivalence.

The second PD workshop took place on March 14, 2011, and teachers learned about different techniques designed to teach mathematical equivalence to students. During the second workshop, teachers were introduced to more formal instruction techniques on equivalence and the equal sign. After the workshop, teachers used two of these techniques with their students, one technique per classroom, over a two- to three-week period. The third PD workshop took place on May 12, 2011, and teachers collaborated with researchers as part of the larger project to package classroom activities that will be distributed to other teachers in Canada.

Participants

The participants of this study were third- and fourth-grade students from four classrooms in three different elementary schools in the Montreal area. There were two grade 3 classrooms, one grade 4 classroom, and one grade 4/5 classroom. Fifty-six students ($N = 56$) participated in the study. There were 11 students from one of the grade 3 classes, 15 students from the other, 19 students from the grade 4 class, and 11 fourth-graders from the grade 4/5 class. The participants had an average age of 9-10 years (9-2 for third-graders and 10-2 for fourth-graders).

The participants' parents signed a consent form allowing their children to take part in the study. The students also gave their personal consent before all testing.

Design

The design of the present study in the context of the larger project is presented in Figure 1. The participants completed paper-and-pencil tests (the Equivalence Test) before and after instruction. The Equivalence Test, which measured students' baseline understanding of equivalence concepts, was given less than a week before the participants' teachers took part in the first PD workshop. During the second PD

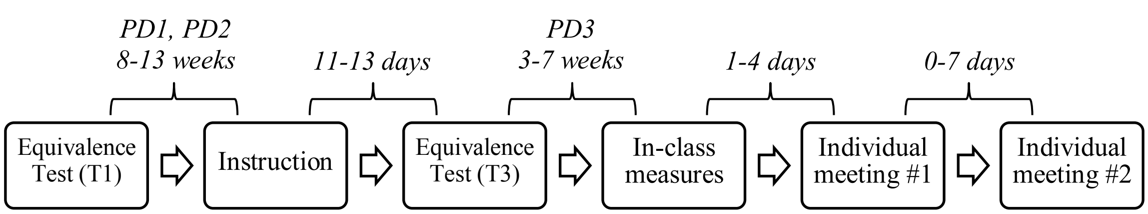


Figure 1. Study design.

workshop, the teachers were given information about two different types of instruction: one based on inquiry methods and another based on principles of direct instruction (called here the "direct approach"). The teachers implemented the inquiry instruction in one of their classrooms and the direct approach in another. The students in the sample were those who received the direct approach only. Thus, students from four classrooms participated in the present study.

During the second PD, the teachers watched a demonstration on how to use the direct approach in their classroom, and they also engaged in a role-playing activity to familiarize themselves with the content of a lesson plan. The lesson required teachers to first explicitly explain to their students that the goal of solving an equivalence problem (e.g., 3 + 1 + 1 = 3 + __) is, "to find a number that fits in the blank so that when you put together the numbers on the left side of the equal sign, you'll have the same amount as when you put together the numbers on the right side of the equal sign." Following this explanation, the teachers solved a series of equivalence problems on the board in front of their students while emphasizing that both sides of the equal sign need to represent the same amount. At the end of the lesson, teachers were asked to give their students some practice problems that they could do in small groups or on their own. To ensure treatment fidelity, the teachers were observed during their instruction. Researchers on the project video-recorded all classroom lessons, which were reviewed by the project team with respect to the sequencing and manner in which the teachers covered the equivalence principles with their students.

An isomorphic version of the Equivalence Test was given to students before the third PD workshop, eleven to thirteen days after their teachers had used the formal instruction techniques in their classrooms (April 2011). Three to seven weeks after that, students were given two additional tests in their classroom: the Fluency Test (Watchorn & Bisanz, 2005), designed to measure students' automaticity of addition and subtraction facts, and the Rating Definitions Task (McNeil & Alibali, 2005a; McNeil & Alibali, 2005b; Rittle-Johnson & Alibali, 1999; Watchorn, 2011), one measure of students' conceptual understanding of the equal sign.

The first of two individual meetings were conducted with each student one to four days later. During the first meeting, students were given: the Test of Nonverbal Intelligence 3rd Edition (TONI-3; Brown, Sherbenou, & Johnsen, 1997), the WJ-III Numbers Reversed (Woodcock, McGrew, & Mather, 2001), and the WJ-III Auditory Working Memory (Woodcock et al., 2001). The second meeting took place about one week after the first. During the second meeting, which was video-recorded, students completed the Symbolic Task and the Nonsymbolic Task (Sherman & Bisanz, 2009). The results of these additional tests helped me gain some insight on the variables that can explain the observed differences in the students' performance on equivalence problems.

Instruments and Measures

Equivalence Test. As a measure of their understanding of mathematical equivalence, all participants were asked to complete the Equivalence Test (created by Watchorn & Bisanz, 2005). Two isomorphic versions of this paper-and-pencil test were administered to students before and after instruction.[1] Each test consisted of 29 problems, and students were asked to solve as many as they could by writing down their answer on a blank line provided in the equation. Students first solved five practice addition and subtraction problems followed by 20 symbolically presented equivalence

---

[1] Because the same test was administered three times for the larger project, for counterbalancing purposes, three isomorphic versions of the test were used. Only the first and third administrations were considered for the present study.

problems (e.g., 6 + 7 = ___ + 5; see Appendix A). There was one canonical addition or subtraction problem presented between each group of five noncanonical equivalence problems. There were four types of noncanonical equivalence problems in this measure: (a) identity, (b) commutativity, (c) two-term part-whole, (d) three-term part-whole, and (e) combination (Sherman, 2004). The blank was immediately following the equal sign in half of the equivalence problems and was at the end of the equation in the other half.

Students received 1 point for each correct answer and 0 points for each incorrect answer. The score for each item was added to obtain a total score. Students' total scores were converted to percent. Separate scores were calculated for canonical and noncanonical problems.

Additional in-class measures. After the Equivalence Test posttest, children were given two additional measures in class: the Fluency Test (designed by Watchorn & Bisanz, 2005) and the Rating Definitions Task (McNeil & Alibali, 2005a; McNeil & Alibali, 2005b; Rittle-Johnson & Alibali, 1999; Watchorn, 2011).

Fluency Test. The Fluency Test (Watchorn & Bisanz, 2005) is meant to assess children's automaticity of basic addition and subtraction facts. Watchorn and Bisanz used this measure in their study for the same purpose. The test is a paper-and-pencil measure consisting of 39 addition and subtraction problems. Students are asked to answer as many as they can in a given amount of time by writing down their answers on blank lines provided in the equations. Third-graders were given 105 seconds and fourth-graders were given 90 seconds to complete the test.

Correct responses were given a score of 1, and incorrect responses were given a score of 0. Students' score on the Fluency Test was calculated by dividing the number of correct responses by the amount of time students were given to complete the task. The minimum score possible was 0, and the maximum score possible was 22.3 for third-graders and 26.0 for fourth-graders.

Rating Definitions Task. In the Rating Definitions Task (McNeil & Alibali, 2005a; McNeil & Alibali, 2005b; Rittle-Johnson & Alibali, 1999; Watchorn, 2011), students are asked to rate six different definitions of the equal sign by circling one of three faces (unhappy, neutral, happy) corresponding to their evaluation of the definition ("not so smart," "kind of smart," "very smart"; see Appendix B for all six items). Two definitions reflect a relational understanding of the equal sign (e.g., "both sides of the equal sign should have the same amount"), two reflect an operational understanding of the equal sign (e.g., "the answer goes next"), and two are not related to equivalence (e.g., "all the numbers after it are small").

Students were awarded 2 points for "very smart" ratings, 1 point for "kind of smart" ratings, and 0 points for "not so smart" ratings for each one of the two relational definitions. For the operational and unrelated definitions, students were awarded 2 points for "not so smart" ratings, 1 point for "kind of smart" ratings, and 0 points for "very smart." Students' scores on the task consisted of the total of the points they had received across the six definitions and ranged from 0 to 12. Students with higher scores supported a more relational definition of the equal sign.

Individual meetings. Together with the additional in-class measures, the additional tests and questions asked during the individual meetings served to identify the student characteristics that could explain the variation in performance on equivalence problems. More specifically, the meetings included tests that corresponded to the hypothesized predictors of performance on the posttest measure.

Four research assistants in the larger project conducted the individual meetings. Each student took part in two individual meetings. During the first session, students were administered the TONI-3 (Brown et al., 1997) as a measure of their nonverbal intelligence, and the WJ-III Numbers Reversed (Woodcock et al., 2001) and WJ-III Auditory Working Memory (Woodcock et al., 2001) as two measures of their working memory. During the second session, students' conceptual understanding of the equal sign in a symbolic context was assessed through the Symbolic Task, specifically through children's justifications of their answers and the Generating Definitions component of the Symbolic Task. Students' conceptual understanding of the equal sign in a nonsymbolic context was assessed through the Nonsymbolic Task (Sherman & Bisanz, 2009).

TONI-3. The TONI-3 (Brown et al., 1997) is a nonverbal measure of cognitive ability that requires students to solve abstract/figural problems. The instructions are given nonverbally, with gestures and facial expressions, to each student and five practice items are administered prior to the administration of the test.

Students are presented with items printed in a Picture Book. There are 45 items in total, and each one is printed on a single page; items are arranged from easiest to most difficult. Each item requires the student to choose, from several response choices, the picture that best fits in the empty box of a stimulus pattern. Students are asked to respond by pointing to the answer they believe is the correct one. After the practice items, test administration begins with the first item and ends once the student makes three incorrect responses in five consecutive items ("ceiling") or once all 45 items are administered.

Each student's raw score was calculated by adding the number of correct responses between Item 1 and the ceiling item. The TONI-3 manual contains tables that were used to translate the raw scores into standardized scores. The standardized scores were used in the analyses.

The authors reported a high degree of reliability (Brown et al., 1997) for the TONI-3. More specifically, they found that the internal consistency of the test is .93, the test-retest reliability is .91 for one form of the test and .92 for the other, and that the interscorer reliability is .99. In terms of construct validity, results from a study by Banks and Franzen (2010) indicated that the TONI-3 is positively correlated with the Wechsler Intelligence Scale for Children – Fourth Edition's Full Scale IQ; more specifically, the authors report a strong positive correlation with the Matrix Reasoning subset. Also, Brown et al. (1997) found that the TONI-3 scores were correlated to measures of school achievement.

WJ-III Numbers Reversed. The WJ-III Numbers Reversed is a measure of working memory (Woodcock et al., 2001). The examiner states a string of random digits between 0 and 9, and students are asked to repeat the numbers in reverse order. The strings of digits become increasingly longer as the test progresses. The administration of this test ends once a child makes three or more errors in a block of items or once all items have been administered.

Students received 1 point for each string of numbers they correctly repeated backwards. The points earned were added to obtain a total score, which could range from 0 to 30. Woodcock et al. (2001) reported that the reliability of the WJ-III Numbers Reversed for 8-year-olds is .86 (as cited in Seethaler & Fuchs, 2006).

WJ-III Auditory Working Memory. The WJ-III Auditory Working Memory is another measure of working memory (Woodcock et al., 2001). The examiner states a string of random digits between 0 and 9 and words (e.g., 3, bread, 1, lion), and students are asked to repeat the words first and then the digits in the order in which they are said. The strings of digits and words become increasingly longer as the test progresses. The administration of this test ends once a child answers all three items of a block of items incorrectly or once all items have been administered.

Students received 2 points if they repeated the words in the correct order followed by the digits in the correct order. They received 1 point if they attempted the words first and repeated either the words or the digits in the correct order. If they were not able to repeat the words or the digits in the correct order or if they started with the digits, they received a score of 0. The points earned were added to obtain a total score, which ranged from 0 to 42. The median test reliability of the WJ-III Auditory Working Memory for 2-19 year-olds is .88 (Schrank, 2011).

Symbolic Task. In the first part of the Symbolic Task, students were asked to solve five equivalence problems presented symbolically (similar to the problems that were on the Equivalence Test; 3 + 4 = 4 + ___, 6 + 9 = ___ + 9, 6 + 4 = 5 + ___, 4 + 5 + 3 = ___ + 6, and 6 + 4 + 3 = 6 + ___). Each problem was presented on an individual index card; students wrote their answers directly on the card. They received a score of 1 for every correct answer and a score of 0 for every incorrect answer. Scores on each of the five items were added to obtain a total score that could range from 0 to 5. In addition, students were asked to give a justification to the interviewer for how they solved the problem. Specifically, the students were asked, “Can you tell me why you wrote down

___?" The second part of the task required students to generate their own definition of the equal sign when asked what the "=" symbol meant in the equation 3 + 4 = 2 + 5 presented on an index card. The interviewer pointed to the equal sign and asked, "What is the name of this symbol? Can you explain to me what this symbol means?"

Coding of justifications. The justifications given by students in the first part of the Symbolic Task were coded, regardless of whether the answer was right or wrong, as either (a) Relational, (b) Procedural, (c) Operational, or (d) Other. Justifications were coded as Relational when students mentioned that the left side and the right side of the equal sign needed to represent the same amount. For example, as a response to the equation 3 + 4 = 4 + __, a Relational justification would be, "the left side has to equal the same as the right side. The left side is 7, and if I put a 3 on the line, the right side is going to equal 7 too." Justifications were coded as Procedural if students described a procedure for solving the equation without explaining the rationale behind the procedure. An example of a Procedural justification when solving the equation 6 + 4 + 3 = 6 + ___ would be, "First you add 6 plus 4 plus 3, that's 13. Then you do 13 minus 6, and that's 7." When students offered explanations that reflected misconceptions such as "the answer comes next" or "add all the numbers," these justifications were coded as Operational. For example, when solving the problem 6 + 4 = 5 + ___, an Operational justification would be, "The answer is 15 because you have to add up all these numbers: 6 plus 4 plus 5." Justifications were coded as Other when they did not correspond to any of the previously described categories. An example of an Other justification would be, "I just put a random number on the line."

Coding of generated definitions. In the second part of the Symbolic Task, students were asked to generate their own definitions of the equal sign. Definitions were coded as either (a) Relational, (b) Operational, or (c) Combined. The definitions were coded as Relational when students explained that both sides of the equal sign needed to represent the same amount. An example of a Relational explanation would be, "The equal sign means that we need to have the same thing on each side." Definitions were coded as Operational when descriptions contained a misconception, such as "add all the numbers" or "the answer comes next." An example of an Operational definition would be, "The equal sign means to add the two numbers and to put the answer of those two numbers next to it." Definitions were coded as Combined when students provided both relational and operational aspects in their definitions. An example of a definition in this category would be, "The equal sign always means different things. Sometimes it means the same thing, like the same answer [relational], and sometimes it means you need to say the answer of the problem [operational]."

Nonsymbolic Task. To assess students' conceptual understanding of the equal sign in nonsymbolic contexts, they were asked to solve equivalence problems using manipulatives (Sherman & Bisanz, 2009). The different problems were represented using small wooden cubic blocks placed on white index cards (see Appendix C). There were two (or three, depending on the item) index cards on the left side of a piece of cardboard folded like a tent, and two index cards on the right side; the blocks on each index card represented the amounts in the arithmetic expressions. To make the two sides of the folded cardboard piece distinct, the index cards on the left side were placed on red construction paper, and the index cards on the right side were placed on green construction paper.

To familiarize children with the task, the materials were arranged to represent 1 + 1 = ___. The interviewer then solved the problem by placing two wooden blocks on the right index card and asked children if the answer was correct. Once the children understood the task, they were presented with five arithmetic problems. There was one problem for each of the following types: (a) identity (i.e., 5 + 6 = ___ + 6), (b) commutativity (i.e., 4 + 5 = 5 + ___), (c) two-term part-whole (i.e., 5 + 3 = ___ + 2), (d) three-term part-whole (i.e., 4 + 3 + 5 = 2 + ___), and (e) combination (i.e., 4 + 3 + 5 = ___ + 4). Children were instructed to "put blocks on the empty [card][2] so that when you put together these on this side of the blue tent [interviewer pointed to the blocks on the left], you'll have the same number as when you put together these on this side of the blue tent" (interviewer pointed to the blocks and the empty card; Sherman & Bisanz, 2009, p. 91). Children were asked to justify their answers.

Students were awarded 1 point for each correct answer and 0 points for each incorrect answer. A total score was calculated by adding all correct answers. The maximum total score on this task was 5, and the minimum total score was 0.

Procedure

Equivalence Test. The third- and fourth-grade students of the four participating teachers were given three isomorphic versions of the Equivalence Test. In each one of the four classrooms, approximately one third of the participants received version A of the test at the pretest, approximately one third received version B, and approximately one

---

[2] Sherman and Bisanz (2009) used plastic bins instead of index cards in their study.

third received version C. As part of the larger study, students were given a test between the pretest and the posttest. In each classroom, approximately half of the students who had received version A at pretest were given version B, and the other half were given version C. Approximately half of the students who had received version B at pretest were given version A, and the other half were given version C. Approximately half of the students who had received version C at the pretest were given version A, and the other half were given version B. For the posttest, each student received a version of the test (A, B, or C) that was different from the ones they had received for the pretest and for the second test.

The tests were given to the students in their intact classrooms. Each test was administered during one of the students' regularly scheduled mathematics class and lasted approximately 30 minutes. The students stayed seated at their desks and worked individually on the test. I clearly read the test's instructions and explained them to the students before they began.

Additional in-class measures. I went into each classroom to administer the Fluency Test and the Rating Definitions Task to the students. This testing session took place three to seven weeks after the posttest during one of the students' regularly scheduled mathematics class and lasted approximately 15 minutes. The students were given a worksheet that had the Rating Definitions Task on one side and the Fluency Test on the other side. They were seated at their desk and asked to work individually on the test. They were also asked to wait for my signal before starting each task or turning over the page. I explained the Rating Definitions Task first, and I read each definition out loud twice. I gave students a few seconds to circle their answer before moving on to the next item. I then explained the Fluency Test and asked students to stop writing when I would tell them that their time was up.

Individual meetings. Project research assistants met students individually on two occasions and held the first session one to four days after the in-class measures had been administered. The second meeting took place a maximum of one week after the first. For each one of these individual meetings, the student was taken out of his or her classroom for a period of approximately 45 minutes. The meetings took place in a quiet room in the school. The first meeting was devoted to the TONI-3, the WJ-III Numbers Reversed, and the WJ-III Auditory Working Memory. The second session was devoted to the Symbolic Task and the Nonsymbolic Task. The second was video-recorded, and the focus of the camera was only on the child’s hands.

Chapter 4: Results

Descriptive Statistics

Performance on noncanonical problems at T3. Across both grades 3 and 4, students' mean performance on noncanonical problems was 26.91% (*SD* = 38.01) at pretest (T1) and 70.58% (*SD* = 38.47) at posttest (T3) (see Table 3). The distribution of scores at T1 and T3 are presented in the boxplots in Figure 2. While the T1 and T3 medians are very different (5% at T1; 95% at T3), it is interesting to note that the bottom quartile at T3 was below 40% (*n* = 14), meaning that a quarter of the students were still unable to solve eight or more noncanonical problems correctly out of 20 after instruction. When students answered problems incorrectly, they either used the "use all the numbers" strategy, the "answer comes next" strategy, or used a strategy that could not be deciphered.

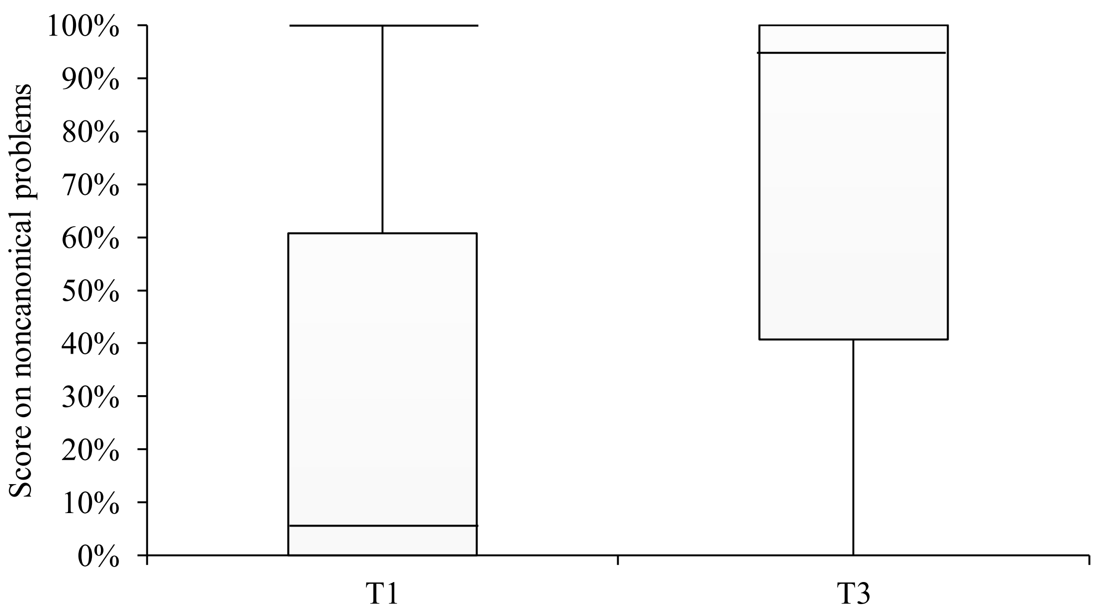


Figure 2. Distribution of T1 and T3 performance scores on noncanonical problems.

Intercorrelations among measures. Intercorrelations between all variables are presented in Table 1. Age was not significantly correlated with any of the variables. The percentage of noncanonical problems answered correctly at T3 (the outcome measure) was significantly correlated with all variables ($p < .05$) except Rating Definitions Task scores.

Effects of time, grade, and equation type. Means and standard deviations for performance on canonical and noncanonical problems at T1 and T3 as a function of grade are presented in Table 2. A 2 x 2 x 2 analysis of variance (ANOVA) was conducted. There were two within-group factors: time (T1, T3) and equation type (canonical, noncanonical); and grade was the between-group factor (grade 3, grade 4). No significant main effects or interactions were found with grade as a factor ($p > .05$). There was a main effect of time, $F(1, 54) = 65.82$, $p < .001$, indicating that all students improved from T1 to T3 regardless of grade level or equation type. There was a main effect of equation type, $F(1, 54) = 107.60$, $p < .001$, indicating that performance on canonical problems was higher than performance on noncanonical problems. There was an interaction between time and equation type, $F(1, 54) = 50.81$, $p < .001$ indicating that time had different effects on students' performance depending on equation type.

Simple effects analyses indicated that performance on canonical problems ($M = 86.26$, $SD = 17.40$) was significantly higher than performance on noncanonical problems ($M = 26.91$, $SD = 38.01$, $p < .001$) at T1. Similar results were found at T3, again with performance on canonical problems ($M = 90.08$, $SD = 16.27$) significantly higher than on noncanonical problems ($M = 70.58$, $SD = 38.46$, $p < .001$). Finally, simple effects analyses revealed a significant improvement on noncanonical problems

Table 1

Intercorrelations Among Variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	–												
T1 Measures													
2. Noncan. T1	-.07	–											
3. Can. T1	.18	.33	–										
T3 Measures													
4. Noncan. T3	-.06	.47	.49	–									
5. Can. T3	.16	.29	.58	.49	–								
In-class Measures													
6. Fluency Test	.14	.45	.37	.48	.36	–							
7. Rating Definitions	.05	.15	.09	.10	.15	.25	–						
Individual Meeting Measures													
8. TONI-3	-.12	.15	.43	.53	.44	.19	.13	–					
9. Numbers Reversed	.01	.29	.52	.50	.44	.46	-.07	.49	–				
10. Auditory WM	.09	.32	.54	.51	.57	.44	.05	.58	.77	–			
11. Symbolic	.18	.43	.37	.66	.43	.54	.17	.33	.49	.44	–		
12. Nonsymbolic	.17	.32	.37	.42	.49	.22	.03	.34	.29	.33	.45	–	
13. Generating Def.	.25	.35	.32	.40	.30	.32	.33	.42	.31	.35	.45	.34	–

Note. Correlations greater than .29 are significant at the .05 level, two-tailed; correlations greater than .35 are significant at the .01 level, two-tailed; and correlations greater than .47 are significant at the .001 level, two-tailed.

from T1 (*M* = 26.91, *SD* = 38.01) to T3 (*M* = 70.58, *SD* = 38.47, $p < .001$). No significant improvement was observed for canonical problems ($p = .064$).

Table 2

Means (in %) and Standard Deviations by Grade

	Grade 3 (n = 26)		Grade 4 (n = 30)	
Measures	*M*	*SD*	*M*	*SD*
Canonical Problems (T1)	81.94	21.36	90.00	12.24
Canonical Problems (T3)	87.61	17.86	92.22	14.71
Noncanonical Problems (T1)	28.98	34.35	25.12	41.43
Noncanonical Problems (T3)	69.89	39.50	71.17	38.22

Because there were no significant effects found when grade was a factor, and because age was not significantly correlated with any of the measures, all data for grade 3 and grade 4 students were combined in the subsequent analyses. The means and standard deviations of the scores on all measures across grades can be found in Table 3.

Table 3

Means and Standard Deviations for All Measures at T1, T3, In-class, and Individual Meetings

Measures	*M*	*SD*	*n*
T1 Measures			
Noncanonical Problems (T1)	26.91[a]	38.01	56
Canonical Problems (T1)	86.26[a]	17.40	56
T3 Measures			
Noncanonical Problems (T3)	70.58[a]	38.47	56
Canonical Problems (T3)	90.08[a]	16.27	56
In-class Measures			
Fluency Test[g]	6.42[b]	4.32	55
Rating Definitions Task	6.57[c]	1.87	54
Individual Meeting Measures			
TONI-3	101.09	11.36	53
Numbers Reversed	10.55[d]	3.58	55
Auditory Working Memory	18.69[e]	6.32	54
Symbolic Task	70.36[a]	44.15	56
Nonsymbolic Task	90.37[a]	21.19	54
Generating Definitions Task	.84[f]	.78	56

[a] Reported in percent [b] max: 20, min: 0 [c] max: 12, min: 0
[d] max: 30, min: 0 [e] max: 42, min: 0 [f] max: 2, min: 0
[g] To compare third- and fourth-graders' scores, a new maximum value was calculated. No grade 3 student completed more than 35 items, and no grade 4 student completed more than 30 items. We used these maximum numbers of items to calculate the new maximum score of 20 (i.e., grade 3: 35 items/1.75 minutes, grade 4: 30 items/1.5 minutes).

Mean performance on noncanonical problems at T3 (M = 70.58) and on the Symbolic Task (M = 70.36) was almost identical, which suggests that the students maintained performance from T3 to the second individual meeting (during which the Symbolic Task was administered), which was held three to eight weeks later. At T3, performance on the Nonsymbolic Task (M = 90.37) was significantly higher than performance on the Symbolic Task (M = 70.36), $t(53) = -3.342$, $p < .01$. Furthermore, the variance of scores on the Symbolic Task was higher than the variance of scores on the Nonsymbolic Task. Together, these data show that students performed better on a task assessing their understanding of equivalence (i.e., Nonsymbolic Task) than on a task assessing their understanding of the equal sign symbol (i.e., Symbolic Task) at T3.

Predictors of Performance on Noncanonical Equivalence Problems

Four hierarchical regression analyses were conducted to explain the variance in students' performance on noncanonical problems at T3, as assessed through the Equivalence Test. The same three predictor variables were entered in all four analyses, and they were as follows: performance on noncanonical problems at T1, TONI-3 standard scores, and Fluency Test scores. Performance on noncanonical problems at T1 was used to control for students' prior knowledge about the equal sign, the TONI-3 was used as a control for general ability, and the Fluency Test was used to control for automaticity of basic addition and subtraction facts. The fourth predictor variable was different for each of the four analyses: (1) Numbers Reversed scores, (2) Generating Definitions Task scores, (3) Rating Definitions Task scores, and (4) Nonsymbolic Task scores. The Numbers Reversed assessed a domain-general factor (i.e., working memory), and the other three tasks assessed factors specific to mathematical equivalence. Only the

Number Reversed measure was used – and not the Auditory Working Memory measure – because the two measures were highly correlated ($r = .77, p < .001$).

Table 4 presents the results of the first hierarchical regression analysis. This analysis revealed that performance on noncanonical problems at T1 predicted 21% of the variance in performance on noncanonical problems at T3, TONI-3 scores accounted for an additional 23% of the variance, and Fluency Test scores accounted for an additional 5%[3]. Scores on the Numbers Reversed task, however, were not found to significantly predict the performance on the outcome measure after the first three variables had been accounted for.

Table 4

Hierarchical Regression Analysis Summary for Variables Predicting Performance on Noncanonical Problems at T3 (N =51)

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.21	
Noncanonical problems T1[a]	0.25	0.12	.26*		
Step 2:				.44	.23***
TONI-3[b]	1.36	0.39	.42**		
Step 3:				.49	.05*
Fluency Test[c]	2.02	1.08	.24		
Step 4:				.50	.00
Numbers Reversed[d]	0.75	1.41	.07		

$^*p < .05$ $^{**}p < .01$ $^{***}p < .001$

[a] $df = 1, 49$ [b] $df = 1, 48$ [c] $df = 1, 47$ [d] $df = 1, 46$

[3] The sample sizes were slightly different for each analysis, which explains why the variance for the first three predictor variables is not exactly the same in each regression.

In the second hierarchical regression analysis (see Table 5), using the same first three measures as in the previous analysis, Generating Definitions Task scores did not significantly explain the variance in performance on noncanonical problems at T3 once the first three predictor variables had been taken into account.

Table 5

Hierarchical Regression Analysis Summary for Variables Predicting Performance on Noncanonical Problems at T3 (N =52)

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.22	
Noncanonical problems T1[a]	0.26	0.12	.26*		
Step 2:				.43	.21***
TONI-3[b]	1.39	0.38	.42**		
Step 3:				.49	.06*
Fluency Test[c]	2.30	1.06	.27*		
Step 4:				.49	.00
Generating Definitions Task[d]	2.23	6.00	.05		

$^*p < .05$ $^{**}p < .01$ $^{***}p < .001$

[a]$df = 1, 50$ [b]$df = 1, 49$ [c]$df = 1, 48$ [d]$df = 1, 47$

The third hierarchical regression analysis (see Table 6) revealed that Rating Definitions Task scores did not predict the variance in performance on the outcome measure once the first three variables had been accounted for.

Table 6

*Hierarchical Regression Analysis Summary for Variables Predicting Performance on Noncanonical Problems at T3 (*N *=50)*

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.21	
Noncanonical problems T1[a]	0.27	0.12	.27*		
Step 2:				.46	.25***
TONI-3[b]	1.74	0.38	.49***		
Step 3:				.51	.06*
Fluency Test[c]	2.40	1.04	.28*		
Step 4:				.52	.00
Rating Definitions Task[d]	-1.04	2.22	-.05		

*$p < .05$ ***$p < .001$

[a]$df = 1, 48$ [b]$df = 1, 47$ [c]$df = 1, 46$ [d]$df = 1, 45$

In the last hierarchical regression analysis performed (see Table 7), scores on the Nonsymbolic Task did not account for the variance in the outcome measure after the first three predictors had been taken into account.

Table 7

*Hierarchical Regression Analysis Summary for Variables Predicting Performance on Noncanonical Problems at T3 (*N *=50)*

Step and predictor variable	*B*	*SE B*	β	R^2	ΔR^2
Step 1:				.21	
Noncanonical problems T1[a]	0.26	0.12	.26*		
Step 2:				.43	.23***
TONI-3[b]	1.64	0.42	.44***		
Step 3:				.50	.06*
Fluency Test[c]	2.37	1.02	.28*		
Step 4:				.50	.00
Nonsymbolic Task[d]	0.09	0.23	.04		

$*p < .05$ $***p < .001$

[a]$df = 1, 48$ [b]$df = 1, 47$ [c]$df = 1, 46$ [d]$df = 1, 45$

Relationship between Equal Sign Definitions and Performance on the Symbolic Task

Symbolic Task justifications. Definitions of the equal sign were first measured using the students' justifications for their answers. A 2 x 3 chi-square analysis, with item as the unit of analysis (n = 273), was conducted to test for a relationship between scores obtained on Symbolic Task items (i.e., 0 or 1) and justifications (i.e., Relational, Operational, Procedural) given for each item (see Table 8). The "Other" justification was removed from the analysis because it occurred infrequently, resulting in expected counts less than 5.

Table 8

Frequencies of Scores on Symbolic Task Items by Justification Type

Justification Type	Item Score	
	0	1
Relational	0 (0%)	87 (44.6%)
Procedural	2 (2.6%)	108 (55.4%)
Operational	76 (97.4%)	0 (0%)
Total	78 (100%)	195 (100%)

Results indicated that item score was significantly related to justification type, $\chi^2(2) = 263.38$, $p < .001$. More specifically, 76 of 78 items answered incorrectly were associated with an Operational justification, and all of the items answered correctly were associated with either a Relational (n = 87, 44.6%) or a Procedural justification (n = 108, 55.4%). The high proportion of correctly solved items associated with a Procedural justification can be explained by the fact that 73.1% of these items (79/108) were attributed to students who gave a Relational justification in at least one other item on the task, suggesting that they held a relational understanding of the equal sign but did not verbalize it in their justifications for all five items.

To test the relationship between justification type and performance on the Symbolic Task using student as the unit of analysis required grouping the students according to justification type. Two groups were formed based on the justifications the students gave to the five items: (1) Relational: when students gave at least one relational justification and no operational justification during the task; and (2) Operational: when students gave at least one operational justification and no relational justification during the task. Seven students did not fit in either group and were removed from the analysis. One student gave both Relational and Operational justifications during the task, and six students were removed because they only gave a Procedural explanation without any justification (Relational, Operational, or Other). An independent-samples *t*-test was then performed to compare the mean scores on the Symbolic Task between the Relational group ($n = 33$, $M = 98.79$, $SD = 4.85$), and the Operational group ($n = 16$, $M = 2.50$, $SD = 10.00$), and a significant difference was found, $t(47) = 45.67$, $p < .001$.

Generating Definitions Task. A second way of assessing the students' definitions was to ask them directly what the meaning of the equal sign symbol was. Using this measure, students generated three definition types in different proportions: 23.2% of students ($n = 13$) generated a relational definition, 39.3% ($n = 22$) generated an operational definition, and 37.5% ($n = 21$) generated a combined definition. Means and standard deviations for performance on the Symbolic Task and on noncanonical problems at T3 by definition type are presented in Table 9.

Table 9

Performance (in %) on Symbolic Task and on Noncanonical Problems at T3 by Type of Definition Generated

Measures	Relational (n = 13) *M*	*SD*	Operational (n = 22) *M*	*SD*	Combined (n = 21) *M*	*SD*
Symbolic Task	92.31	27.74	44.55	47.78	83.81	35.56
Noncanonical Problems T3	85.00	31.75	49.00	42.06	84.25	27.64

Two one-way ANOVAs were performed with definition type as the independent variable and performance on the Symbolic Task and performance on noncanonical problems at T3 as the dependent variables in each analysis. A significant effect of definition type on performance on the Symbolic Task was found, $F(2, 53) = 7.94, p < .01$. Post-hoc comparisons conducted between each pair of means with a Bonferroni correction revealed no significant difference between the Relational group ($M = 92.31$, $SE = 7.69$) and the Combined group ($M = 83.81$, $SE = 7.76$, $p = .99$), but each outperformed the Operational group ($M = 44.55$, $SE = 10.19$, $p = .003$ for the Relational comparison and $p = .006$ for the Combined comparison).

There was also a significant effect of definition type on performance on noncanonical problems at T3, $F(2, 53) = 6.93, p < .01$. Again, post-hoc comparisons conducted with a Bonferroni correction between each pair of means revealed no significant difference between the Relational group ($M = 85.00$, $SE = 5.14$) and the

Combined group (M = 84.25, SE = 6.03, p = .99), but each outperformed the Operational group (M = 49.01, SE = 8.97, p = .014 for the Relational comparison and p = .005 for the Combined comparison).

Chapter 5: Discussion

The present study aimed to determine the factors that could explain children's individual differences in performance on equivalence problems. More specifically, the links between students' performance and their domain-general abilities (i.e., working memory) and domain-specific abilities (i.e., conceptual understanding of the equal sign in symbolic and nonsymbolic contexts and ability to generate definitions of the equal sign) were examined.

Third- and fourth-graders were asked to solve a series of equivalence problems before receiving instruction on the subject, and again after instruction. To explain students' differences in performance after instruction on these problems, several additional measures addressing possible predictors were administered. After instruction, students completed the Fluency Test and a measure requiring them to evaluate definitions of the equal sign. They were then met individually and were administered the TONI-3 (as a measure of their general ability) and two working memory measures (i.e., Numbers Reversed and Auditory Working Memory). During these individual meetings, students were also given equivalence problems that were presented symbolically and nonsymbolically, and they were asked to explain how they solved these problems. In addition, students were asked to generate their own definition of the equal sign.

As expected, scores on canonical problems were higher than scores on noncanonical problems both before and after instruction. Furthermore, although performance on noncanonical problems significantly improved after instruction, there was still considerable variance in the performance on the noncanonical problems, indicating that instruction was not enough for many students (see also Watchorn et al., 2011). The additional measures were used to explain the variance in performance on noncanonical problems after instruction.

Four hierarchical regressions were conducted, but none of the Numbers Reversed scores, Generating Definitions Task scores, Rating Definitions Task scores, or Nonsymbolic Task scores were able to account for the variance in the outcome measure's scores after the first three predictor variables (i.e., performance on noncanonical problems at T1, TONI-3 scores, Fluency Test scores) had been taken into account. There are several possible explanations for these results. Many studies support the claim that working memory is related to mathematical performance (e.g., Alloway & Passolunghi, 2011; Fuchs et al., 2010), which it why it was surprising to find that it did not explain the variance in performance in the present study. Raghubar et al. (2010) recognized the link between working memory and mathematics, but explained that the relationship between the two is complex and that other factors need to be taken into consideration, such as the types of mathematical skill and the particular working memory component involved (see also Rasmussen & Bisanz, 2005). In their investigation of working memory components related to the mathematics achievement of second- and third-graders, Meyer et al. (2010) found that the central executive and phonological loop components of working memory predicted the mathematics performance of second-graders, but not of third-graders, and that the only component that predicted third-graders' performance was the visuo-spatial sketchpad. Interestingly, Meyer et al. found that their "Backward Digit Recall" measure – a central executive measure similar to the Numbers Reversed measure used in the present study – did not predict mathematics performance for second- or third-grade students. I might have found significant results in terms of working memory's contribution to children's performance on equivalence problems had I used a measure targeting a different aspect of working memory (e.g., a measure assessing the visuo-sketchpad), but research in this area is inconclusive (see also De Smedt et al., 2009).

Students' ability to generate equal sign definitions also did not explain the variance in performance on noncanonical problems. In McNeil and Alibali's (2005a) study, participants generated different definitions of the equal sign depending on the context in which the symbol was presented. This suggests that the Generating Definitions Task alone may not give an accurate indication of the students' complete conceptions and knowledge of the equal sign, which may explain the non-significant result in the regression analyses. Along the same lines, Seo and Ginsburg (2003) found that children can have relational and operational views of the equal sign at the same time, but that the context in which the equal sign is presented will impact the view they will hold. For example, a student's performance on equivalence problems might reflect a relational understanding of the equal sign, but that same student may offer a definition of the symbol more closely related to an operational view. As Rittle-Johnson, Matthews, Taylor, and McEldoon (2011) explain, it may be more difficult for students to give verbal definitions than to solve written problems. Different skills might be required for generating definitions and for applying the concepts they convey.

Although the Rating Definitions Task has been used in previous research as a measure of conceptual understanding (e.g., Rittle-Johnson & Alibali, 1999), the results of this study suggest that it is not, by itself, as valid a measure as others have believed. The same phenomenon as with the Generating Definitions Task might be occurring: it may be more difficult for children to evaluate definitions than to apply equivalence concepts when solving problems. In an effort to reduce the time children spent outside their classroom for the individual meetings, the Rating Definitions Task was administered in a whole-class setting. Researchers who have used this measure in previous work (e.g., McNeil & Alibali, 2000, 2005b) administered it to each participant individually. It is possible that the task may not be as valid when used in a group context as opposed to an individual context, and that had I administered it in a different way, I would have found significant results.

As hypothesized, Nonsymbolic Task scores did not explain the variance in performance on noncanonical problems. Furthermore, a majority of students reached ceiling for the Nonsymbolic Task, and there was a low amount of variance in the data on this measure. That performance on this task was significantly better than performance on the Symbolic Task suggests that for children, there might be a disconnect between concepts and symbols. This result is consistent with Sherman and Bisanz's (2009) finding that children are able to perform well on equivalence problems presented nonsymbolically (i.e., with concrete materials), but that they are not always able to map this understanding to problems of the same type presented symbolically. It is also possible that the teachers in this study did not put enough emphasis on the connection between concepts and symbols during instruction. Previous research has suggested that instruction focusing on such connections is related to students' conceptual understanding and mathematical performance (e.g., Matthews & Rittle-Johnson, 2009; Rittle-Johnson & Alibali, 1999).

In an attempt to better understand the individual differences in students' conceptual understanding of equivalence, transcripts of children's justifications of how they solved problems on the Symbolic Task and of their generated equal sign definitions were coded and then analyzed. Consistent with previous research (Knuth et al., 2006; McNeil & Alibali, 2005b; Sáenz-Ludlow & Walgamuth, 1998), the analyses indicated that they generated Relational and Operational definitions, which aligns with the literature. Furthermore, children who generated only operational definitions of the equal sign performed poorly on noncanonical problems, which is also consistent with previous research. Students who generated Relational and Combined definitions of the equal sign performed equally well on noncanonical problems, suggesting that if students have some element of relational understanding of the equal sign in their conception of the symbol, they are able to solve noncanonical equations correctly.

In the present study, several students generated Combined definitions, which is something that is not discussed in much detail in the literature. Rittle-Johnson et al. (2011) had asked participants to explain what the equal sign meant. They mentioned that some of the children who initially gave an operational or ambiguous definition of the equal sign gave a relational definition after being probed (i.e., "Can it mean anything else?"). The authors did not, however, further discuss this occurrence in their article. Knuth et al. (2006) also found that some students gave both relational and operational elements in their definition of the equal sign, but they decided to code these responses as "relational" in their analysis of the data.

An operational view of the equal sign develops by default and is not easy to overcome (Rittle-Johnson et al., 2011). The children who gave a Combined definition might be in a transitional stage taking them from a strictly operational view of the equal sign to a relational view. Watchorn (2011) provided some evidence to this effect. Her study revealed that some of her second- and fourth-grade participants also held combined views of the equal sign.

Periods of conceptual change have been associated with high levels of variability in children's strategies; the process of change involves the addition of new structures, followed by the removal or integration of old structures (Alibali, 1999). Participants in the present study who held a combined view of the equal sign might have been in a period of conceptual transition, which would explain their use of a dual representation (i.e., relational and operational) to explain the meaning of the equal sign. Since within-child variability has been associated with subsequent learning (Siegler, 2007), children from the Combined group might be on their way to developing an understanding of the equal sign as a solely relational symbol. Students who hold combined views may have understandings that are split into different "microworlds" (Lawler, 1981), and it might only be a question of time and accrual of experiences before their dual view of the equal sign transforms into a stable relational definition.

Limitations and Future Research

There were some weaknesses to the present study. First of all, the size of the sample ($N = 56$) was small; a larger sample would have yielded more power to detect significant results for the regressions. Also, most measures were only administered at the second time point (i.e., after instruction), and therefore it is difficult to determine if children's performance on these measures, particularly the domain-specific ones, was due to instruction or to pre-existing characteristics. Because performance on the Symbolic Task and on the Equivalence Test at T3 was almost identical, and because the items on both measures were very similar, I presume that had the Symbolic Task been administered at T1, its scores would have been similar to those of the Equivalence Test at T1. Unfortunately, there are no data that we can use to make assumptions about the performance on the other conceptual measures before instruction.

Furthermore, for some students, a considerable time (up to ten weeks) elapsed between the instruction and the second individual meeting in which the Symbolic Task was administered, which could impact the amount of conceptual information the students retained. Lastly, all analyses were correlational, and as such they cannot reveal causation. A possible follow-up experiment would involve randomly assigning students to groups who would receive identical instruction on the equal sign from the same teacher. The students would be assessed before and after instruction on their working memory abilities, ability to solve equivalence problems, conceptual understanding of the equal sign in symbolic and nonsymbolic contexts, and ability to generate definitions of the equal sign. Such an experiment would address some of the limitations of the current study; specifically, changes in performance on the different measures could be attributed to instruction or to students' personal pre-existing characteristics, and the development of conceptual understanding could be more clearly understood.

I believe that further research is also necessary to explain why students were generating combined definitions of the equal sign even though they were able to successfully solve equivalence problems. A study evaluating the effects of instruction type (conceptual, procedural, or combined) on students' knowledge in the context of solving equivalence problems could be conducted. Students could be interviewed about the strategies they use while solving equivalence problems before, during, and after instruction on the topic. This could shed light on the trajectory of students' understanding in terms of the instructional features they tend to focus on and the interpretation they make of the instruction they are receiving.

As discussed above, none of the three domain-specific measures used in the regression analyses predicted students' performance on the outcome measure. Rittle-Johnson et al.'s (2011) study aimed to address what they see as a problem that exists in the mathematical equivalence research field. The authors deplore the paucity of evidence for the reliability and validity of the measures typically used to assess mathematical equivalence, and believe there needs to be a standard measure of equivalence knowledge. The possibility that currently-used measures lack validity and reliability might explain, perhaps, why no significant results were found in the present study with the three measures of mathematical equivalence understanding.

I also found that the individual interviews afforded the opportunity to gain a more complete picture of how students were really thinking about the equal sign and its meaning. I was able to detect subtleties in their explanations that a paper-and-pencil measure would not have afforded. I think that more sensitive ways to investigate students' conceptions of the equal sign (e.g., generating definitions and justifying answers) might be important interview strategies to use in future research on the topic.

Implications

There are several educational implications that can be derived from the present study. First, in accordance with the literature, students have the ability to understand equivalence, but struggle with problems involving the equal sign symbol. Instruction should therefore focus on explicitly linking symbols to the concepts they represent.

Teachers should also use caution when using paper-and-pencil tests as a sole indicator of students' understanding of the equivalence concept, even if they are able to solve equivalence problems correctly. The findings of the current study suggest that teachers may find it beneficial to engage in discussions with their students because some of them might have an incomplete conceptual understanding of the equal sign. If understanding is equated with performance, having a Combined definition of the equal sign is not a problem. I believe, however, that a complete understanding of the meaning of the equal sign should also be reflected in how children conceptualize and define equivalence. It is important to ensure that children eventually resolve their fragmented understanding so that they will avoid facing difficulties when tackling more complex problems upon starting high school algebra classes (Hammer, 2010; McNeil & Alibali, 2005a). Future research examining the long-term effects of having a partial understanding of the equal sign is needed to better understand its consequences on future performance in mathematics.

References

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences, 21*, 133-137. doi:10.1016/j.lindif.2010.09.013

Banks, S. H., & Franzen, M. D. (2010). Concurrent validity of the TONI-3. *Journal of Psychoeducational Assessment, 28*, 70-79. doi:10.1177/0734282909336935

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

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). Test of Nonverbal Intelligence–Third Edition. Austin, TX: PRO-ED.

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development, 19*, 81-93. doi:10.1016/j.cogdev.2003.10.001

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

Carpenter, T. P., & Levi, L. (2000). Developing conceptions of algebraic reasoning in the primary grades. Research report prepared for the National Center for Improving Student Learning and Achievement in Mathematics and Science. Madison, WI.

Carraher, D. W., Schilemann, A. D., Brizuela, B. M., & Earnest, D. (2006). Arithmetic and algebra in early mathematics instruction. *Journal for Research in Mathematics Education, 37*, 87-115.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology, 103*, 186-201. doi:10.1016/j.jecp.2009.01.004

Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics, 6*(4), 232-236.

Franke, M. L., Carpenter, T. P., & Battey, D. (2008). Content matters: Algebraic reasoning in teacher professional development. In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 333-359). New York, NY: Lawrence Erlbaum Associates.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., … Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities. *Developmental Psychology, 46*(6), 1731-1746. doi:10.1037/a0020662

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*, 1-16. doi:10.1002/acp.934

Gilmore, C. K., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetical skill. *British Journal of Educational Psychology, 76*, 309-331. doi:10.1348/000709905X39125

Hammer, K. (2010, August 31). Students' early struggles with math count later on, study shows. *The Globe and Mail*. Retrieved from http://www.theglobeandmail.com

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339-366. doi: 10.1080/01443410500341056

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*, 82-88. doi:10.1016/j.lindif.2009.07.004

Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5-17). New York, NY: Lawrence Erlbaum Associates.

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

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

Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction, 19*, 513-526. doi:10.1016/j.learninstruc.2008.10.002

Lawler, R. W. (1981). The progressive construction of mind. *Cognitive Science, 5*, 1-30.

LeFevre, J., DeStefano, D., Coleman, B., & Shanahan, T. (2005). Mathematical cognition and working memory. In J. I. D. Campbell (Ed.), *The handbook of mathematical cognition* (pp. 361-378). New York, NY: Psychology Press.

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

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. doi:10.1016/j.jecp.2008.08.004

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466. doi:10.1016/j.cogsci.2003.11.002

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

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

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. C., 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*(3), 367-385.

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., & Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences, 20*, 101-109. doi:10.1016/j.lindif.2009.08.004

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110-122. doi:10.1016/j.lindif.2009.10.005

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91,* 137-157. doi:10.1016/j.jecp.2005.01.004

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

Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *British Journal of Educational Psychology, 00*, 1-19. doi:10.1348/000709908X398106

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

Sáenz-Ludlow, A., & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics, 35,* 153-187.

Schrank, F. A. (2011). Woodcock-Johnson III tests of cognitive abilities. In A. S. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 415-434). New York, NY: Springer.

Seethaler, P. M., & Fuchs, L. S. (2006). The cognitive correlates of computational estimation skills among third-grade students. *Learning Disabilities Research & Practice, 21*(4), 233-243.

Seo, K., & 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: Constructing adaptive expertise* (pp. 161-187). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Sherman, J. (2004). *Constraints on understanding arithmetic equivalence: Effects of misunderstanding the equal sign*. Unpublished manuscript.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*, 104-109. doi: 10.1111/j.1467-7687.2007.00571.x

Stephens, A. C. (2006). Equivalence and relational thinking: Preservice elementary teachers' awareness of opportunities and misconceptions. *Journal of Mathematics Teacher Education, 9*, 249-278. doi:10.1007/s10857-006-9000-1

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

Watchorn, R. (2011). *Improving Children's Understanding of Mathematical Equivalence* (Doctoral dissertation). Retrieved from ProQuest. (NR73300)

Watchorn, R., & Bisanz, J. (2005, October). Enduring instructional effects on equivalence problems. Poster presented at the meeting of the Cognitive Development Society, San Diego, CA.

Watchorn, R., Osana, H. P., Sherman, J., Taha, M., & Bisanz, J. (2011, October). Improving mathematical equivalence performance in Grades 2 and 4. Poster proposed to the Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Reno, NV.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Itasca, IL: Riverside Publishing.

Appendix A

Sample Equivalence Test

6 + 4 = ___

12 – 2 = ___

12 + 3 - 6 = ___

11 - 6 = ___

7 + 6 - 4 = ___

3 + 4 = 3 + ___

3 + 5 + 4 = 2 + ___

6 + 4 = 4 + ___

7 + 5 + 3 = 7 + ___

6 + 9 = ___ + 7

4 + 7 = ___

4 + 5 = ___ + 3

3 + 6 + 4 = ___ + 3

7 + 8 = 8 + ___

4 + 5 + 6 = ___ + 2

4 + 7 = ___ + 7

9 + 3 – 4 = ___

6 + 2 = 2 + ___

6 + 8 = ___ + 8

7 + 3 = 4 + ___

6 + 4 + 3 = 5 + ___

5 + 6 + 4 = ___ + 5

9 – 3 = ___

5 + 3 + 4 = 5 + ___

5 + 7 + 3 = ___ + 4

9 + 5 = 6 + ___

5 + 3 = ___ + 3

5 + 6 = ___ + 5

7 + 3 = ___

Appendix B

Rating Definitions Task

$$\mathbf{3 + 4 + 1 = 3 + \underline{\quad\quad}}$$

		Very smart	Kind of smart	Not so smart
1	Billy said that "=" means "the answer to the problem"	☺	😐	☹

		Very smart	Kind of smart	Not so smart
2	Kim said that "=" means "the end of the problem"	☺	😐	☹

		Very smart	Kind of smart	Not so smart
3	Marc said that "=" means "that two amounts are the same"	☺	😐	☹

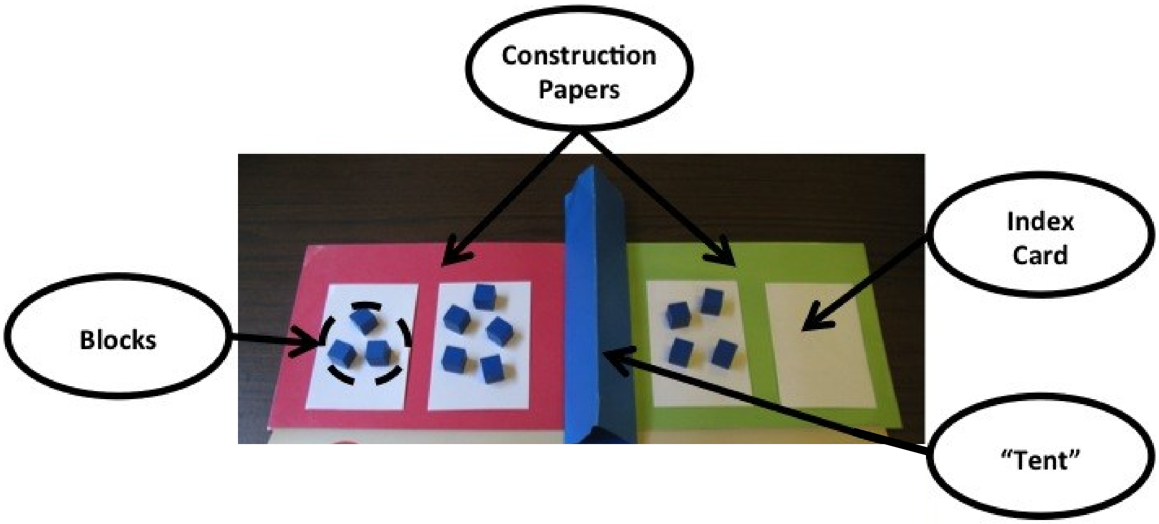
		Very smart	Kind of smart	Not so smart
4	Lisa said that "=" means "to repeat the numbers"	☺	😐	☹

		Very smart	Kind of smart	Not so smart
5	John said that "=" means "that something is equal to another thing"	☺	😐	☹

		Very smart	Kind of smart	Not so smart
6	Michelle said that "=" means "the total"	☺	😐	☹

Appendix C

Nonsymbolic Task



Verbal instructions:

I want you to put cubes on the empty card so that when you put together these [point to the cubes on the left of the tent] on this side of the tent, you'll have the same number as when you put together these [point to the cubes and empty card] on this side of the tent.