MOVING ON A PATH VERSUS COLLECTING OBJECTS: NEW PERSPECTIVES TO ANALYZE STUDENT LEARNING WITH INTEGER MODELS

By

Julie Nurnberger-Haag

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Educational Psychology and Educational Technology—Doctor of Philosophy

2015

ABSTRACT

MOVING ON A PATH VERSUS COLLECTING OBJECTS: NEW PERSPECTIVES TO ANALYZE STUDENT LEARNING WITH INTEGER MODELS

By

Julie Nurnberger-Haag

Most mathematics beyond middle school requires that students operate with all real numbers, both positive and negative quantities. Later mathematics (e.g., linear and quadratic equations, coordinate graphing, absolute value equations, transformations, and matrices) and science topics (e.g., vectors, sound waves, and charged particles) all require use of negative numbers. Yet, some of the identified difficulties students have with more advanced studies, such as algebra, are due to difficulties with negative number arithmetic and notation (Vlassis, 2008).

With the intent to help students overcome these difficulties, multiple instructional models are promoted in school textbooks and teacher resources. Multiple models are used although little is known about what students learn with these models. Different integer instructional models have different implications for learning mathematics, because they draw on different conceptual metaphors and students physically move their bodies in different ways to enact these models. In light of conceptual metaphor theory, the difficulties students have had with negative numbers and even those of mathematicians in history (Hefendehl-Hebeker, 1991), reveal that the collecting objects metaphor may be a cognitive obstacle to those first learning negative number arithmetic. Moreover, consistency of humans' physical motions with targeted ideas is a factor of cognition, so the influence of students' physical movements on their learning may be a critical, yet underexplored factor.

In order to compare how these model differences influence learning, this study randomly assigned eight classes of initial learners to a specific collecting-objects (chip model) or moving-on-a-path metaphor-based model (number line model) to learn integer arithmetic with the four primary operations (addition, subtraction, multiplication, and division). This dissertation presents the main results of this pre-post-delayed posttest study to answer the questions: (a) With respect to ordering numbers and integer arithmetic, what do students demonstrate learning by enacting each model and what, if any differences in learning are found between models? (b) What meanings do students express for "-" symbols (negative signs, subtraction signs, or opposite signs) and how does the integer model used influence these meanings? Findings indicated that either instructional model did support significant learning gains for integer arithmetic and qualitative expression of basic "-" symbol meanings. These findings, moreover, did support theory that a motion-aligned model using a moving-on-a-path metaphor (walk-it-off number line model) was a better first model, because it supported initial integer learning better than a collecting objects metaphor based model both for integer arithmetic learning and opposite meanings of "-."

To Eric, my partner in life for the last 25 years

ACKNOWLEDGEMENTS

I have much for which to be thankful in terms of the support from friends and family. My graduate experience was such that the mathematics education faculty and graduate students made this a place that seeks to bring out the best of each of us. I am grateful for the colleagues, mentors, mentees, and friends who have become part of my life. I was able to find the research assistance, school sites, and funding necessary to do this work, just when I needed it. I am grateful for each of these new timely opportunities and the people who made these opportunities possible.

My Committee

To my dissertation committee, Drs. Jack Smith (Chair), Kristen Bieda, Sandra Crespo, and Raven McCrory, thank you for being the kinds of mentors I knew I needed when I was choosing a graduate program. Each of you epitomize what educators should do—assist people on the path they chart for themselves. You are people that students can count on to be genuine regardless of the context in which they work with you. I looked forward to every meeting I had with you individually and as a group and left feeling empowered to go forward. Each of you played significant roles in my graduate experience at different times during my program prior to and during my dissertation. I look forward to continuing the relationships we started here.

Jack, thank you for helping me determine if the EPET program was the best fit among the many ways a person could focus on mathematics education at MSU, your advocacy on my behalf to obtain fellowships and an assistant professor position, your excellent feedback and your friendship. The way you provide feedback so that critique is constructively encouraging even when the need for it (due to the difficulty of scholarly activities) is inherently discouraging, will serve as a model for mentoring my own students.

Kristen, thank you for first introducing me to the theories of Lakoff and Nunez before I even applied to Michigan State University and being a critical sounding board for my different research projects. Most of all, thank you for being so intentional about mentoring graduate students about early career work and becoming my friend.

Raven and Sandra, I appreciated your enthusiasm for my research and bringing me as a guest into your undergraduate and graduate classes, which helped me to stay connected with the reasons I do this work during years when I did not teach courses myself. I appreciate how both of you call it like you see it and always wanted the best for me. I look forward to the many ways and reasons we will find for our paths to cross in the future.

To Claire Vallotton, thank you for developing the course HDFS 924. I promote your course with anyone in education who will use video data or observe classrooms. I believe the timing of the opportunity to take your course was crucial to implementing the qualitative aspects of this work for my dissertation and my future research. Your ability to mentor students from both theoretical and practical standpoints is a model for teaching graduate students.

To Ralph Putnam, although you were not on my committee, I wanted to formally acknowledge that I appreciated your time, effort, and guidance through MTHE 954 as I started writing the proposal and for funding some of the expenses in your role as an EPET Program Coordinator.

To the administrators, teachers and students who participated in the study whom I do not name to protect their identities, I am grateful to you for allowing me to share your precious educational time. I look forward to future collaborations.

Resources

I gratefully acknowledge funding that made this study possible. First, a Michigan State

University Distinguished Fellowship made it possible throughout the first five years of my program to focus more time on my research than many students are afforded. Second, Michigan State University Dissertation Expense fellowships from the College of Education and the Educational Psychology & Educational Technology Program made it possible to conduct the study. Most recently, I appreciate The National Academy of Education and National Academy of Education/Spencer Dissertation Fellowship Program that funded the analysis and writing of this dissertation. Moreover, I value the mentoring and peer collaborations the National Academy of Education/Spencer Dissertation Fellowship Program provided.

Additional offices at Michigan State University that served a role in my education as well as this dissertation study I also acknowledge here in the hope that these important resources continue to recruit and serve MSU graduate students: the Institute for Research on Teaching and Learning (IRTL) in the College of Education, Center for Statistical Training & Consulting (CSTAT), College of Education Service Center, and the Center for Teaching and Technology (TIES) (specifically Tracy Abbott, Megan Drangstveit, David Reyes-Gastulem, Emre Gonulates, and Terri Gustafson).

Data Collection and Analysis

Thank you to Liz Pauly. You were a great help on a different research project, and when I asked you if you could find me someone else as wonderful as you to assist with my dissertation research you introduced me to Hope Younggren.

Hope, I am grateful to you for volunteering, your enthusiasm, your careful consideration of student thinking, attention to detail, and among many other things, helping me make some tight turn around times analyzing pretests in order to be ready to interview and teach the students. I look forward to staying in touch as you become a teacher.

Thanks also to Andrew Krause. You helped make it possible to conduct enough post interviews to implement my study before the students had yet another Michigan snow day off school before their winter break.

Elizabeth Wood I am grateful to you for your attention to detail, hours of data entry, and help with improving some of the written text. I look forward to continuing to work with you as you determine how you will use your love of writing to determine a career path.

Last, but certainly not least Erin Novelly, I am so excited for you as you begin your teaching career. Thank you for your attention to detail, commitment to understanding how students think and learn, and thoughtful debates to ensure that much of the qualitative coding of this dissertation could help us uncover information that might benefit future students and teachers. I look forward to continuing to stay in touch as your mentor and friend.

Family and Friends

To my partner in life, my husband Eric Haag. As I wrote this dedication I realized that I am completing this milestone 25 years after our first date. Although everyone who completes a dissertation has at least one significant supporter who sacrificed in some ways while the dissertator did this work, you are a rare and amazing partner. In high school even before we started dating, you were my cheerleader. Each year from that time you have done more and keep getting better to help me be who I want to be. Commuting three hours a day so that I could be close to campus was certainly more than most spouses would be willing to do, and I do not take that for granted. When I see my name in print before I hyphenated it with yours, it just looks incomplete and not quite me, just as I would be without you. You help me be who I am and to do everything I want to do in this world, so I am glad publicly to brag about you and thank you.

To my Mom Hazle and Eric's Mom Jan, I know you both prefer private expressions of

appreciation, so I will not say a lot here, but I needed to publicly say what amazing Moms you both are. You both epitomize the unconditional love and support that mothers should have for their children. I am so grateful that the right graduate program for me was MSU, which allowed me to move home to the Lansing area to spend more time with you when things were going well and have your support when things were rough.

To my Dad, Dr. Fred Nurnberger, the fact that you earned a Ph.D. yourself meant that since childhood earning such a degree always seemed possible. You always told me that I could be anything I wanted to be and thought about how to make it happen. For example, I remember in elementary school I told you I wanted to be an astronomer, so you told me I would need to major in astrophysics and you got out the Michigan State University catalog to show me the kinds of classes I would take. It is a nice connection for us to have both earned our doctorates from Michigan State University, 43 years apart.

To Pete and Deborah, thanks for rooting for me during the challenging times and helping me to celebrate successes. Most of all, Pete as a die-hard University of Michigan fan, thanks for wearing your MSU dad cap I bought you my first year in the program and being a good sport when I make you drink out of the MSU coffee mug.

Kari, thanks for being my sister and someone outside my field who can commiserate about the challenges and understand the joys of a research-related career path. I would also like to mention Aunt Pat, who was always passionate about learning and research, but who is no longer with us. I miss you and the conversations we would have had about my research had you been with us.

Sara Soderstrom, words don’t do justice to the gratitude I have for the fate that brought us together in Russia 20 years ago this year and the knowledge that we’ll have each other’s support for the next 40 or more years. Thank you Amy Scheuermann, Maryl Randel, Mariana Levin, and Jamie Wernet for becoming such close friends and critical readers of my practicum and other work that helped me as I continue to develop as a scholar. Sam Otten, your efforts as the student member of my practicum committee, discussions of my work at MLRG and conferences, advice as an early career faculty member, and your friendship are much appreciated. Thank you Eryn Stehr, Rani Satyam, and the Mathematics Learning Research Group for engaging conversations that helped me improve my work and develop friendships. I look forward to staying in touch throughout our careers.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

CCSS-M	Common Core State Standards for Mathematics
CMP	Connected Mathematics Project
CMT	Conceptual Metaphor Theory
EDT	Explain and Draw Test
I	Interviewer
I/A	Interviewer/Author
IAT	Integer Arithmetic Test
MD	Multiplication and Division
TIMSS	Trends in International Mathematics and Science Study

Overview and Introduction to Dissertation

Humans are always moving, and the ways we move influence how and what we think (Antle, 2011; Barsalou, 2008; Lakoff & Nunez, 2000). Students' physical motions in learning environments deserve greater recognition as an ever-present and potentially critical factor of learning. Many of these motions that likely influence individual's thinking in any given moment are idiosyncratic to a person in a particular situation. Such idiosyncratic motions might warrant research. With the intent for greatest impact on mathematics learning and teaching, however, I choose to focus on those physical movements that affect thousands of students at a time due to instructional practices. Model-movements in particular deserve attention because they are sets of motions that regardless of whether educators and researchers attend to these movements, the instructional approaches used encourage certain patterns of students' physical motions.

My research draws attention to model-movements used in mathematics instruction as one way to assess the educational impact of students' physical motions during classroom instruction and contribute to the design of better instructional models. Models are not simply manipulatives. I use the term 'models' to mean not only tools or manipulatives, but also the physical and verbal processes used, afforded, constrained, and chosen with those tools. For example, Dienes' blocks, Digi-blocks, and number lines are tools that can be used in varied pedagogical models. Many different models based on these manipulatives are used to teach students number concepts and skills with inconsistent and contradictory results (e.g., Fuson 1990; Kamii, Lewis, & Kirkland, 2001; Star & Nurnberger-Haag 2011). Insights from cognitive science research show that physical movements that are aligned with concepts support general cognition and learning, whereas misaligned movements interfere with cognition (Day & Goldstone, 2011; Glenberg & Kaschak, 2002; Goldin-Meadow, Cook, & Mitchell, 2009). Instructional approaches with certain ways of using tools encourage these patterns of movement to become a model. Thus, *model-movements*, the term I use to describe specific physical motions that each pedagogical model affords, constrains, or encouraged by choice may be a critical and previously overlooked factor of student learning of mathematical ideas.

My research is grounded in my experience as a practitioner. I was troubled by the inconsistencies I noted between the model-movements with manipulatives and the disciplinary conceptions I sought for students to develop about number (whole number multi-digit calculation as well as negative number arithmetic). Early in my career, I observed students' difficulties in making sense of operations with negative number, which led me to design other ways of teaching and researching how best to help initial learners with negative number arithmetic. This dissertation considered these model-movement issues to seek empirically-validated solutions to student learning of negative number arithmetic, because it is an enduring challenge of mathematics learning and teaching that has long-term impact on students' later mathematics performance. To do so I experimentally compared how two integer models impact student learning if these were the first integer models students experience: a typical chip model (a manipulative model), and a particular number line model in which students move their bodies along a number line on the floor (Nurnberger-Haag, 2007).

Dissertation Format

The traditional dissertation format, a book with chapters, has been used in the United States since the 1800s (Berelson, 1960, p.173 cited in Duke & Beck, 1999). This book format, even just decades ago, may have been a useful dissertation format when it was common for new scholars to publish books about research. In this century, however, multiple peer-reviewed articles are required to advance in the tenure process and such publications are likely to impact the field of research more readily than a book. Consequently, this dissertation uses a two-article format in order to better reflect the publication work that contemporary educational researchers in academia do (Duke & Beck, 1999). The corpus of this dissertation centers on the main results of this mixed methods experimental study.

Focus of Each Article

The first article sheds light on what students learned with a particular integer model and relative learning effects for different aspects of integer knowledge based on both short-term assessments and delayed posttests. In order to compare benefits and challenges of particular models, an experiment was needed. The analysis of the experiment in this manuscript is QUANT + qual (Creswell, 2008), meaning that quantitative analysis is emphasized but qualitative aspects of student reasoning were incorporated into understanding individual student's knowledge before statistically comparing the average learning of students who learned with each model. Both overall performance and performance on subconstructs were analyzed. In order to provide a comprehensive analysis of the written assessments across all aspects of integer knowledge tested, a more detailed analysis of each construct will be left for future analysis to be reported in other manuscripts (e.g., a focus on multiplication and division operations).

Whereas the first article provides comprehensive results from the written assessments, the second article uses a qualitative approach to extend prior work on *negativity* (Gallardo & Rojano, 1994; Sfard, 2000; Vlassis, 2002, 2004, 2008) to investigate students' meanings of negative and subtraction signs expressed in interviews after particular instructional approaches. Prior research contributed information about how algebra students, who had in prior years learned about negative number arithmetic, think about these signs. These students were also in Mexico, Israel, and Belgium. So this dissertation study specifically contributes information from a new population of fifth and sixth grade students in the United States. Moreover, this study contributes information about students' meanings of these symbols immediately after initial integer arithmetic instruction and if there are qualitative differences due to the model used, given that instruction in the United States often uses many models.

Structure of Dissertation

In the sections that follow, each article appears in succession with tables and figures appearing in Appendix A numbered consecutively from Article One to Article Two as dissertation formatting required. As is typical of a dissertation, appendices B and C include information my committee requested.

ARTICLE ONE: TO WALK THE PATH OR COLLECT THE CHIPS: THE IMPACT OF METAPHORS AND MOTIONS ON LEARNING INTEGER ARITHMETIC

Abstract

Different integer instructional models have different implications for learning mathematics, because they draw on different conceptual metaphors and students physically move their bodies in different ways to enact these models. In order to compare how these model differences influence learning, this study randomly assigned eight classes of initial learners to a specific collecting-objects (chip model) or moving-on-a-path metaphor-based model (number line model) to learn integer arithmetic with the four primary operations (addition, subtraction, multiplication, and division). This pre-post-delayed posttest study addressed the question: What do students demonstrate learning with each model and what, if any differences in learning are found between models? Written test results that showed that both models supported student learning, but the number line model fostered greater learning gains of each aspect of integer knowledge assessed. These findings support the claim that a motion-aligned model drawing on a moving-on-a-path metaphor supported initial integer learning better than a collecting objects metaphor-based model.

To Walk The Path Or Collect The Chips: The Impact Of Metaphors And Motions On Learning Integer Arithmetic

Most mathematics beyond middle school requires that students operate with all real numbers, both positive and negative quantities. Later mathematics (e.g., linear and quadratic equations, coordinate graphing, absolute value equations, transformations, and matrices) and science topics (e.g., vectors, sound waves, and charged particles) all require use of negative numbers. Yet, some of the identified difficulties students have with more advanced studies, such as algebra, are due to difficulties with negative number arithmetic and notation (Vlassis, 2008). In order to prepare students for these later demands, students in upper elementary and middle school need to find answers to problems involving negative quantities as well as understand negative quantities as valid numbers (Ball, 1993; Vlassis, 2008). Students around the world, however, find this knowledge difficult to develop (Altiparmak & Ozdogan, 2010; Gallardo, 2002; Kilhamn, 2011; Perisamy & Zaman, 2011; Pierson Bishop, et al., 2014; Ryan & Williams, 2007; Vlassis, 2008; Warfield & Meier, 2007).

To address these difficulties, I designed and implemented an experimental comparison of two integer models to investigate if either might be more cognitively ergonomic to support learning (for a related discussion of cognitively ergonomic see Artigue 2002 & Abrahamson 2009). Just as we should assess how people use physical tools for how physically ergonomic these processes are with the human body, to improve mathematics instruction we need to assess how cognitively ergonomic models are for learning specific topics. This study contributes to resolving two enduring challenges in mathematics education: one practical and one theoretical. The first concerns improving the way that classroom-based research can inform teachers' practical decisions about using models to teach integer arithmetic. The second offers new insights into the theoretical and practical debate about whether and how physical experience supports cognitively ergonomic mathematics learning.

Research on Integer Understanding and Model-Based Instruction

To situate the issues specific to improving integer arithmetic instruction, I first summarize integer knowledge that students need and then instructional approaches used to address these needs, focusing on models used. Finally, I discuss the theoretical perspectives with which I frame student learning with the common integer model types (chip and number line models) in terms of general cognitive processes they likely involve: conceptual metaphors and consistency of physical movements.

Integer Understanding

Some of the ways in which mathematically competent people need to think about negative and positive quantities are in terms of points on a number line, opposite numbers are points on opposite sides of zero on a number line, opposite values sum to zero, ordering signed numbers, and how to perform all four primary operations (Bofferding, 2014; Chiu, 2001; Lakoff & Nunez, 2000; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010; Thompson & Dreyfus, 1988). In the United States, the *Common Core State Standards for Mathematics (CCSS-M)* includes the above goals and specifies that instruction about negative numbers occur in sixth grade with operations on all rational numbers in seventh grade (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). A thorough understanding of arithmetic and algebra with negative numbers requires understanding three meanings of the written symbol "-" as (a) part of a number, (b) a subtraction operation, and (c) an opposite operation (Vlassis, 2004). The meaning of "-" as part of a number or a structure means people need to recognize, for example "-4" itself as a valid number or structure by itself as opposed to needing to think of it as a result of operations such as 6- 10 (Gallardo, 2002; Sfard, 2000). This structural-numeric understanding of the symbol’s meaning is, however, not sufficient. The other two meanings of the “-“ symbol are operations: *subtract* and *opposite* (Vlassis, 2004, 2008). Mathematically competent students need to understand the binary operation of subtraction of negative numbers including all of its meanings: take-away, difference, and motion on a number line (Lakoff & Nunez, 2000; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010; Sfard, 2000, p. 49; Thompson & Dreyfus, 1988). In addition, the unary operational meaning “opposite of” is crucial to student understanding and use in algebra (Nurnberger-Haag, 2007, Vlassis, 2008). For example, in situations such as –(-4), –X, or (-a) + a = 0 neither the subtraction meaning nor negative as a number makes disciplinary sense. Yet, students believe any number with a “-“ sign denotes a negative number, so they often incorrectly conclude that –(-4) can only be negative four (Vlassis, 2008). This operational conception to *take the opposite* is necessary to understand not only when $a<0$ that –a is actually positive number but also the opposite of algebraic expressions, and recognize when the “-“ symbol could have more than one meaning (Thompson & Dreyfus, 1988; Vlassis, 2008).

Some modern students think of zero as an absolute quantity and have difficulty conceiving of negatives as valid numbers similar to brilliant mathematicians in history (Fischbein, 1987; Hefendehl-Hebeker, 1991). For example, when students do multi-digit whole number subtraction and the subtrahend digit is larger than the minuend digit, a common error is to write 0, because they think they cannot take away more things than they have (Brown & Burton, 1978). Consequently, integer operations are difficult. On a standardized test for example, only one-third of fourteen year-old students could answer -6 -3 and less than half could divide -

24 by 6 (Ryan & Williams, 2007, p. 218). Students usually think subtraction makes a number smaller and multiplication larger because of their extensive experience operating on whole numbers (Ryan & Williams, 2007). These generalizations, however, are not valid for all real numbers. A released eighth grade item from the *Trends in International Mathematics and Science Study (TIMSS)* assessed this operation and generalization knowledge: "If n is a negative number, which of these is greatest? (Answer options 3 + n, 3 × n, 3-n, 3 ÷ n)" (International Association for the Evaluation of Educational Achievement, 2005).). Item accuracy rates substantiate that students across the world (40%) found this difficult as well as students in the United States (48%).

Integer Instruction

Integer arithmetic with negative numbers is counterintuitive, yet essential to most mathematics beyond middle school. Negative numbers, notation, and addition and subtraction operations have been extensively studied (Ball, 2003; Gallardo, 2002; Küchemann, 1981; Liebeck, 1990; Linchevski & Williams, 1999; Saxe et al., 2013; Thompson & Dreyfus, 1988; Vlassis, 2008), yet recommendations are contradictory about how to help students adapt their arithmetic concepts to embrace negative numbers (Star & Nurnberger-Haag, 2011). Küchemann (1981) categorized three types of integer instructional methods as (a) cancellation models in which two opposites cancel, (b) number line models, or (c) abstract methods. Although integer thinking and learning has been extensively studied, investigations of integer learning with particular models, particularly those most accessible in classrooms, have yet to be conducted. Despite the absence of research support (e.g. Star & Nurnberger-Haag, 2011; Freudenthal, 1973; Vig, Murray, & Star, 2014), multiple integer models are promoted in methods textbooks for prospective teachers as well as school textbooks, particularly cancellation and certain number line models. These multiple integer models are promoted to use with the same students with an assumption that more models are better, in spite of a dearth of empirical evidence about each individual model or how particular models could work together to support student learning. I focus here on the first two approaches to investigate how student's physical enactments of a single integer arithmetic model influences learning.

Cancellation models. Küchemann (1981) classified cancellation models as "those in which the integers are regarded as discrete entities or objects, constructed in such a way that the positive integers cancel out the negative integers" (p. get page 62-67). Examples include happy and sad faces, hot and cold cubes, charged particles, or different colored objects such as chips or playing cards (Cotter, 1969; Goldin & Shteingold, 2001; Jencks & Peck, 1977; Liebeck, 1990; Ponce, 2007). Instruction with materials that cancel may include context such as scoring games (Liebeck, 1990; Linchevski & Williams, 1999). Some argue cancellation models support student learning of negative number arithmetic, because such models help students' use prior knowledge of positive numbers as objects (French, 2001; Küchemann, 1981; Liebeck, 1990). For example, students can think of subtraction of negative numbers as taking away objects just as students did when they first learned subtraction of whole numbers (French, 2001; Liebeck, 1990). Advocates of these models also claim that they allow students to use the idea of inverse operations and additive inverses by recognizing that one could either take away positives or add in negatives to achieve the same result (Linchevski & Williams, 1999; Semadeni, 1984).

Typical cancellation models, called chip models, use chips of different colors to represent opposite numbers, (e.g., Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006; van de Walle, Karp, & Bay-Williams, 2010). Chips are an instructional representation, whereas a number line is a disciplinary representation. Consequently, some have viewed chips or other cancellation approaches as instructional gimmicks used in school mathematics, but not real mathematics (Roussat, 2010; Umland, 2011). Some have also critiqued cancellation models, because the models require rules that are not rules of the mathematics (e.g., Rousset, 2010; Star & Nurnberger-Haag, 2011; Vig, Murray, & Star, 2014). Recent theoretical analyses of addition and subtraction have critiqued model rules that are not mathematical rules as being points where the models “break” (Star & Nurnberger-Haag, 2011; Vig, Murray, & Star, 2014). For example, some subtraction problem structures violate mathematical rules by requiring students to add additional chips in order to have enough to subtract (Vig, Murray, & Star, 2014). Although Liebeck (1990) concluded from her empirical analysis that the chip model was better than a number line model, she stated that students reported that these problems that required adding extra chips in order to subtract were more difficult. Detailed analysis of these issues with chip models in terms of model-movements will be discussed in the Theoretical Framework section.

Number line models. A particular set of number line procedures frequently appear in on-line resources and textbooks (e.g., University of Chicago School Mathematics Project, 2007; Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006; Math is Fun, 2013), which I call a typical number line model. These resources usually show diagrams of students walking or draw arrows to indicate movements on the number line. In order to distinguish when the sign “-” means the operation subtract as opposed to a negative number, these number line models use different movements on the number line. Specifically, these models typically teach students to think of the signs of the numbers as indicating to face forward (positive) or backward (negative), whereas subtraction signs mean move left on an horizontal number line. Typical number line models for multiplication and division use repeated addition and subtraction using the same forward or backward movement. Student learning of integer addition and subtraction with number line models have been studied in the contexts of turtles who follow rules to move along a number line (Thompson & Dreyfus, 1988), temperatures on a thermometer (Sfard, 2007), elevators going up and down floors (Ball, 1993; Hill, 1968), and representations of world elevation (Sfard, 2007).

Advocates of these models have argued that they avoid the obstacle encountered with cancellation models that treat numbers as actual objects (Freudenthal, 1973). Number line models are also beneficial because they are disciplinary representations used throughout later mathematics (Freudenthal, 1973). Some claim that using number lines is intuitive for addition and subtraction (Freudenthal), whereas others argue that this is true only for addition (Küchemann, 1981). Although Freudenthal advocated inductive-extrapolatory approaches, he claimed if one were to choose a representation to complement student understanding, a number line is the best representation.

Researchers have studied the difficulties and benefits students have when adding and subtracting on a number line (e.g., Bofferding, 2014; Bruno & Martinon, 1999; Ernest, 1985; Thompson & Dreyfus, 1988). Such research involving individual instruction with students from first to sixth grade has reported that students can make progress on such conceptions, but as Thompson & Dreyfus (1988) found, even after 11 individualized lessons on addition and subtraction with a computer turtle moving on a number line, students have difficulty with problems structures that do not have a net effect that uses the sum of the magnitudes. Some have critiqued typical number line models for the rules required to add and subtract within a specific number line representation or the meanings these number line movements attribute to "-" (Bofferding, 2014; Nurnberger-Haag, 2007; Star & Nurnberger-Haag, 2011). Although Hill (1968) used an elevator model using movement between floors as positions on a number line for multiplication of integers, number lines have been infrequently used for multiplication and division.

Limitations

In order to recommend effective instructional practices for this difficult aspect of mathematics, several gaps in the literature need to be addressed. These include lack of research on multiplication and division as an integral aspect of integer arithmetic knowledge, short-term view of instructional impacts, and investigations of a particular model rather than comparison of different model affordances and constraints.

Limited scope of integer knowledge. Research that investigates either addition and subtraction or multiplication and division in isolation may miss critical aspects of student learning and knowledge. Prior integer research, although important, has primarily focused on negative numbers and/or operations of addition and subtraction (Bofferding, 2014; Liebeck, 1990; Linchevski & Williams, 1999; Pierson Bishop, et al., 2014; Saxe et al, 2013; Stephan & Aykuz, 2011; Thompson & Dreyfus, 1988; Vig, Murray, & Star, 2014). Investigations of student learning of multiplication and division are warranted, because recent research on these operations focused on instructional materials (Seidel, 2012) and existent research on student thinking occurred more than 30 years ago (Küchemann, 1981). Moreover, it is crucial to assess student learning of all primary operations in relation to each other to understand how students' might use this knowledge in the future. For once students learn the rule of signs to multiply and divide negative numbers, many students inaccurately use the rule to add and subtract (e.g., two negatives make a plus such as -4 + -5 = 9; French, 2001; Nurnberger-Haag, 2007).

Short-term view. Students need to understand negative number arithmetic to build on and use in complex ways in formal algebra (Vlassis, 2008), so research that investigates longer-term implications of instructional experiences is critical. Yet, most integer research has assessed students' learning during or immediately after instruction (Stephan & Aykuz, 2012; Thompson & Dreyfus, 1981). This short-term view limits the potential practical educational implications, because issues of retention or long-term changes are not addressed (Yelon, Ford, & Golden, 2013). Moreover, recent research suggests that instead of losing knowledge over time as is common, students who learn mathematics or science with physical motions may actually show greater learning gains longer-term (Cook, Mitchel, & Goldin-Meadow, 2008; Hadzigeorgiou, Anastasiou, Konsolas, & Prevezanou, 2009).

Single models investigated. In order for educators to make effective instructional decisions, it is important to compare methods (Nunez, 2012) and to understand how different methods might offer similar or different learning opportunities. Yet, rather than comparing different models to identify model affordances and constraints, most integer instruction research has focused on how students learn with one particular model (Linchevski & Williams, 1999; Stephan & Akyus, 2012), or experimentally compared one model to no model (that some call an "abstract approach") (Moreno & Mayer, 1999). One study did compare a chip model using the context of "scores and forfeits" to an acontextual number line model (Liebeck, 1990). Although this commonly cited study found a chip model to be more beneficial, it did not include a pretest to adequately compare learning gains and the instruction and post-tests differed in many ways between the two groups of 10 students.

Students' physical motions impact learning. Humans' physical movements influence cognition (Antle, 2011; Barsalou, 2008; Glenberg & Kaschak, 2007; Kontra, Fischer, Lyons, & Beilock, 2015). The primary integer instructional models encourage students to physically move objects (chip models) or implicitly move on a number line (even if students are not afforded the opportunity to physically walk along number lines during instruction typical number line models enacted in classrooms at least implicitly suggest physical motions on a number line). Thus, students' model-movements during effortful learning of integer arithmetic likely impact their cognition. Yet, research has not attended to the ways students' physical model-movements alignment with integer arithmetic might influence mathematical achievement.

Theoretical Perspectives

Whereas most references in research refer to "the" number line or chip model as if the tool or representation constitutes a single model (e.g., Liebeck, 1990; Vig, Murray, & Star, 2014), I refer to "a" model. I argue that number lines and chips are representational tools that can be used in multiple ways. Each of the multiple ways that groups of students and teachers use a particular tool (i.e., with some regularity or patterns of processes), constitutes a different model. By the term model, I mean not only the tools, but also the physical and verbal processes used (due to affordances, constraints, and instructional choices) with those tools (Nurnberger-Haag, 2014). The theoretical perspectives I use to analyze learning with integer models will help to illuminate this stance. I frame student learning with the common integer model types (chip and number line models) in terms of general cognitive processes they likely involve: conceptual metaphors and consistency of physical movements. First I discuss how the classes of integer models relate to human cognition in relation to Conceptual Metaphor Theory (CMT). Then within each class of models, based on recent development about the influence of physical body motions on cognition, I discuss how students' physical model-movements are consistent or inconsistent with the mathematics and may or may not be cognitively ergonomic for integer arithmetic.

Conceptual Metaphor Theory

Humans are always moving and, albeit subconsciously, our physical body motions are one way that we develop conceptual metaphors to conceive of abstract ideas, including mathematics (Johnson, 1987; Gentner & Bowdle, 2008). Numbers can be thought of in many ways including metaphorically as quantities of objects, lengths, locations on a path—in cognitive science such metaphorical projections of real-life experiences to understand abstract ideas are called conceptual metaphors (Lakoff & Nunez, 2000). CMT assists with explaining and unifying mathematical thinking to show that and in what ways mathematical thinking is a part of the varied abstract thinking humans do (Lakoff & Nunez, 2000).

Integer understanding in terms of CMT. The conceptual metaphors humans used to develop and continue to use to conceive of integer arithmetic include *object collection*, *motion along a path*, and *measuring stick* metaphors (Lakoff & Nunez, 2000). Our first experiences with the ideas of number are based on collecting and counting real objects, which forms an *object collection* metaphor meaning of number. For centuries, however, the consequence of using this metaphor for number was that Western mathematicians did not believe negative numbers were valid, because having less than no objects was impossible; zero meant nothing or no thing, an absolute quantity (Berlinghoff & Gouvêa, 2002; Hefendehl-Hebeker, 1991; Rotman, 1993). Pascal for example remarked, “I know people who cannot understand that when you subtract four from zero what is left is zero” (as cited in Hefendehl-Hebeker, 1991, p.26 from source in German). Relying on an object collection metaphor in this way interfered with using negative numbers until people began thinking of numbers in ways that developed into number line representations: lengths in different directions using a *measuring stick* metaphor and subsequently numbers as locations, *motion-along-a-path* metaphor (Chiu, 2001; Freudenthal, 1973; Lakoff & Nunez, 2000; Kilhamn, 2011). An object collection metaphor treats numbers as objects with positive or negative numbers being attributes or kinds of numbers that can be grouped into countable quantities. A measuring stick metaphor refers to thinking of numbers as the end of a positive static length in a particular direction, not as positions on a real or imagined path (Lakoff & Nunez, 2000). With a moving-along-a-path metaphor numbers are thought of as points or positions along a path and also as distances to travel on that path (Lakoff & Nunez, 2000). Facility with more than one metaphor may be necessary for expert understanding of negative numbers (Chiu, 2001; Gentner & Bowdle, 2008). Different conceptual metaphors are not simply different approaches, but lead to different consequences for current and later mathematics topics (Gentner & Bowdle, 2008; Nunez, Edwards, & Matos, 1999).

CMT and integer models. Research of integer arithmetic via conceptual metaphors is in its infancy and has focused on what students and adults who had already learned integer arithmetic expressed about these ideas and processes (Nurnberger-Haag, 2013; Chiu, 2001; Kilhamn, 2011), rather than how instructional models that involve metaphor-based physical motions impact what students learn. CMT offers educational researchers valuable insights to identify the ways that different instructional models might afford and constrain students' conceptions of numbers and arithmetic. This theoretical lens differentiates between the mathematical representations and the ways people think with these representations. Küchemann's (1981) definition of *Cancellation models* treat numbers as though they are real objects. By definition then, cancellation models map entirely consistently with an *object collection metaphor*, in which mathematical operations are thought of as moving or collecting quantities of objects (Kilhamn, 2011). Number line representations, in contrast, could be thought of in ways consistent with either a measuring stick or motion-along-a-path metaphor. A measuring stick metaphor refers to thinking of numbers as the end of a positive length in a particular direction (Lakoff & Nunez, 2000). Number line models can also be thought of with a motion-along-a-path metaphor where numbers are treated as locations or points on the path and also directed movements (Kilhamn, 2011; Lakoff & Nunez, 2000; Nurnberger-Haag, 2007), or as a combination of both of these metaphors to be discussed in later studies. Number line models frequently found on-line and in schools do not commonly use the measuring metaphor, so the remaining discussion and study design focus on object collection and motion-along-a-path metaphors. Table 1 describes the mathematical meanings in terms of the conceptual metaphor and physical model-movements for each integer model. In particular it maps each of meanings for numbers, signs of numbers, and operations.

CMT and Model-Movement Relationship with Integers

How people physically move influences how they think and involves metaphorically projecting interactions in the real world to conceive of abstract ideas (Barsalou, 2008; Gentner & Bowdle, 2008). Thus, in order to understand how students think and provide learning experiences that "complement the ways their conceptual systems naturally work" (Nunez, Edwards, & Matos, 1999, p. 62), educational research should explore how conceptual metaphors and students' physical movements encouraged during mathematics instruction are consistent with the content and influence student learning. Conceptual metaphor has been treated as way of thinking or an object of thought that results from physical experiences, which grounds how humans think about abstract ideas (Lakoff & Nunez, 2000). Whereas other research has referred to conceptual metaphors of integer arithmetic similarly using nouns as though these are reified ideas after physical movement (Chiu, 2001; Kilhamn, 2011) or critiqued CMT for this reason (de Freitas & Sinclair, 2014), I use the verb forms (*collecting objects* and *moving-on-a-path*). I use verbs to emphasize that I consider conceptual metaphors to be a label that categorizes particular grounded patterns of interacting with the world as part of an on-going dynamic system (Nurnberger-Haag, 2014). For the individual (re)enacting such patterns physically or only neurologically, this is usually subconscious (Hurtienne, et al., 2010). Table 2 provides detailed examples of how students could operate with negative numbers using each model. Due to space and student difficulties noted in prior research, the table illustrates how students subtract, multiply and take the opposite of numbers. Each of these will be discussed in more detail.

Humans' physical movements that align with ideas support general cognition and learning, whereas inconsistent movements interfere with cognition (Day & Goldstone, 2011; Glenberg & Kaschak, 2002; Goldin-Meadow, Cook, & Mitchell, 2009; Kontra, Lyons, Fischer, & Beilock, 2015). For instance, a set of psychology experiments found that when people moved actual objects away or toward themselves it evoked the ideas "away" or "toward" that benefitted performance on subsequent abstract tasks consistent with participants' prior movements, but interfered when inconsistent (Glenberg & Kaschak, 2002). Goldin-Meadow, Mitchell and Cook (2009) found when individual elementary students were taught to gesture in particular ways to written numerals in non-traditional equality problems that "the more correct their gestures during the lesson, the better children performed on the posttest (2009, p.3). Researchers tested effects of a partially correct gesture by teaching students to use the same physical gesture as the correct gesture condition, but on the wrong numbers, making this physical motion inconsistent with the mathematics. This group whose gestures were inconsistent with the mathematics did worse than the correct gesture group; however, they still did better than the students who were not taught to gesture.

Collecting object model-movements. Using the way I define *model* to not only mean the tools, but also the physical model-movements and language about the tool use, different approaches constitute particular models. Just as there are multiple number line models, there are multiple models that use colored chips. I first discuss how different model-movements with chips constitute different models. Then detail how chip model-movements are consistent or inconsistent with the mathematics.

Different chip model-movements constitute different models. At least two different chip models use sets of colored chips (e.g., black representing positives and red representing negatives; van de Walle, Karp, & Bay-Williams, 2010; Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006), in which students put in chips to represent addition and take out to represent subtraction. As they move chips to represent numbers in this way, the two models use the same language to represent their movements, "put in" or "add" and so forth. To calculate integer multiplication and division with a chip model, students use a repeated groupings meaning of multiplication, which is the way students generally first transitioned from whole number addition to whole number multiplication. Accompanying language might sound like "put in three groups of two negatives." Although using colored chips to represent negative numbers is consistent with a collecting objects metaphor and these objects afford and constrain certain model-movements, instruction encourages students to move in different ways to represent operational processes that constitute different models. With one particular chip model, students always begin a problem by putting in multiple chips to represent a value of zero (e.g., multiple black chips and the same number of red chips). To enact the more common chip model used in U.S. schools, students only add this extraneous zero representation when a problem requires it. For example, using a typical chip model for -3 - -2=-1 described in Table 2 does not require this extraneous addition, but -3 –

2=-5 does (see Table 2).

Chip model-movement inconsistencies with integer operations. Using the theoretical and empirical perspectives from cognitive science about physical motions influence on cognition and CMT, I draw attention to what students' physical model-movements actually model about the mathematical ideas. Thinking about whether model rules are rules of the mathematics (Vig, Murray, & Star, 2014) offers a dichotomous perspective that could map to my approach of considering whether model-movements are consistent or inconsistent with the mathematics. Human activity, however, is continuous. Accordingly, using perspectives about how closely model-movements represent the mathematics, I point out that different problem structures of each primary operation, afford different degrees of consistency between the chip model-movements and the mathematics. The chip model requires students to move this way, because the chips are actual objects.

Addition and subtraction. Addition operations are the most consistent with respect to model-movements, so students may find addition to be the most cognitively ergonomic operation with chips. Model-movement consistency with subtraction, however, varies depending on the problem structure. Some subtraction problems require inconsistent model-movements from the beginning of a problem, others in-process, and other problems involve only consistent model-movements. The subtraction problems discussed in Table 2, for example, represent consistent model-movement (-3 - -2) and inconsistent from the beginning problem structures (-3 – 2). To do 7-10 with chips, however, students can begin the process with seven positives (black chips) and only encounter a model-movement inconsistency during the process, once they have removed seven chips.

Multiplication and division. Students can move in ways consistent with mathematics to multiply one positive and one negative number, because students can use repeated addition of chips where the positive number represents the number of groups of negatives as the example -3 × -2 in Table 1 demonstrates. Student may not find the product of two negatives cognitively ergonomic with this model, however, because students must begin each problem with inconsistent model-movements. A student first has to add in extraneous chips to have enough chips to remove, and then treat negative numbers as taking out that many groups of negatives (see Table 2). Division problems are more nuanced than multiplication with regard to model-movement consistency, because quotients of two negatives can afford more consistent model-movements than problems that involve one positive and one negative. The problem -12 ÷ -3, for example, can be modeled by splitting 12 negatives into groups of three negatives to obtain the answer of 4 groups. Yet, to begin to model 12 ÷ -3, students need to represent zero with multiple chips in order to take away sets of three negatives until the remaining value of the chips is 12. Chip models also are not able to represent the algebraic operation "opposite of."

Although some have emphasized the theoretical affordances of chips models that encourage students to interpret subtraction with negative numbers as a "take-away" operation, chip models are infrequently used for multiplication and division. In spite of the infrequent use of chip models for these operations and the theoretical prediction that the inconsistent model-movements of chips might pose difficulties for student learning, the chip model does offer potential affordances related to meanings of multiplication that the walk-it-off model does not. A chip model reinforces and extends repeated addition meanings of multiplication with whole numbers to negative numbers, unlike the walk-it-off model, which asks students to determine distances to travel by using memorized whole number products or quotients.

Moving on a path model-movements. The typical number line model promoted in curricular resources directs students to move backwards or forwards depending on the sign of a number and to move in a particular direction depending on the operation (i.e., left, right, up, down, positive direction, negative direction). In other words, as Tables 1 and 2 details, these models inform students *which* direction to move. For multiplication and division, these model-movements also run into difficulties due to using repeated addition meanings of multiplication.

The walk-it-off model, in contrast, employs model-movements consistent with integer arithmetic operations taught in school. The written symbols in the walk-it-off model "–" signal whether to change direction by turning the *opposite direction* or in the case of addition or positive values, to maintain direction (See Table 1). Note in the table that the walk-it-off model uses a stable meaning of "-" signs: "move the opposite direction," which work for all primary operation problem structures (see Table 2). The model-movements vary only due the signs of the numbers and operations, not based on problem structures. Addition and subtraction with the walk-it-off model treat the first number of a problem as a starting point or location, then the addition or subtraction sign signals whether to turn the opposite direction and a second number is a directed magnitude, so this indicates whether to turn the opposite direction and how far to move (See Tables 1 and 2). Multiplication and division with the walk-it-off model, as with chip models, begin with zero. Unlike chip or other number line models the walk-it-off model avoids using repeated addition meanings of multiplication and division. Students who are advanced enough to learn multiplication and division with negative numbers should be able to find products and quotients of whole numbers (i.e., the absolute values of the numbers in multiplication or division problems) as detailed about the walk-it-off model in Table 2. This model was also designed to address the opposite operator meaning of the negative sign shown in the last row of Table 2, necessary for algebra, but not afforded by other models (Nurnberger-Haag, 2007; Thompson & Dreyfus, 1988).

As with a chip model, instruction with the walk-it-off model does not need to begin by using explicit instruction. Students can begin exploring addition and subtraction tasks by moving on number lines in ways that make sense to them based on their whole number arithmetic knowledge. Instruction can then help students formalize how they moved and talked about those movements in terms of the meanings used in the walk-it-off model. From these experiences students can then formalize rules or algorithms for integer arithmetic rather than being explicitly told such rules (e.g., multiplying or dividing two negatives results in a positive number, subtracting a number is the same as adding the opposite of the number). Even though such approaches begin with student thinking, some educators who promote using a context with a number line for extended periods of time might still critique such acontextual uses of number lines for requiring students to use arbitrary rules or algorithms (Stephan & Akyuz, 2012). Research that compares such uses of number lines has not yet provided insights about how student learning might differ between acontextual and contextual uses of a number line.

For acontextual ("bare numbers") arithmetic problems typically studied in middle school, the walk-it-off model does not seem to have limitations in terms of model-movement consistency. If, however, students need to solve context-based problems that draw on different metaphors such as measuring length or collecting objects, students who learn with the walk-it-off model might have difficulty thinking about integer arithmetic in these different ways. This study was not designed to investigate such limitations. Measuring Lengths and Collecting Object metaphors are also ways mathematically knowledgeable people can think about integer arithmetic (Chiu, 2001). Although the walk-it-off model uses a number line representation, it uses a moving-on-a-path metaphor, rather than a measuring metaphor. This may mean that even though some students may be able to apply this model to problems involving temperature measured on a thermometer, others may have difficulty since it might not make sense to apply a moving-on-a-path metaphor to comparing static lengths. As in whole number arithmetic when students are able to act out problems that match the problem structures, students found these problems easier (Fennema, Franke, Carpenter, & Carey, 1993). Even if the walk-it-off model were to assist students with using number lines in other contexts in which an actor does not move, ideas of things cancelling that are needed for school and disciplinary science knowledge such as cancellation of electrical charges may be better supported if students learn with a collecting objects metaphor that the chip model could foster.

Focus of Study

A limitation of both chip and walk-it-off models may be that they might make it difficult for students in the future to think of mathematics in disciplinary ways consistent with pure mathematics. For example, integer exponents use a different meaning than either model, pure mathematicians think of the plane moving not the objects on it and conceive of arithmetic with only two operations (addition and multiplication). Such potential limitations relate to claims that integer arithmetic should only be taught in ways that some call abstract methods (e.g., algebraic proofs, arithmetic rules, or noticing patterns of written problems and solutions; Heefer, 2011). Yet, school mathematics at this point does teach students in ways that treat the plane as static and four, rather than two arithmetic operations, thus this study focuses on comparing models and leaves investigation of these longitudinal issues for future research.

A single model may be insufficient to the develop range and depth of integer arithmetic knowledge students need, because each model affords different ways of engaging with mathematically-relevant objects and actions. To effectively use multiple models, however, researchers and teachers need more information about benefits and challenges of each model in isolation and how combinations of models impact student learning. Thus, in order to contribute to the long-term goal of informing how to help students develop a strong sense of negative numbers, related arithmetic, and subsequent algebraic use, this study used an experiment to focus attention on benefits and affordances of a single model with students who had not been influenced by other instructional models

The purpose of this study was to determine which models for integer learning are more beneficial as the first integer model students encounter, for what aspects of integer knowledge, and why, in order to contribute to improved teaching of integer arithmetic. Specifically, the study was designed to contribute to resolving contradictory claims in the field that cancellation and number line models are more intuitive than the other for introducing integer arithmetic. With a goal of practical impact, to test these assertions this study used the theoretical perspectives described to experimentally compare particular cancellation and number line models that could be or are easily implemented in schools and that use model-movements I thought would be the best case of each conceptual metaphor. Potential outcomes could be that the empirical results demonstrate that different models simply promote different ways of thinking about integer arithmetic, but result in similar student performance. Alternatively, the study might provide empirical evidence that using the walk-it-off number line model as the first model (used the first time that students encounter the topic) leads to better student learning for at least some aspects of integer knowledge, because it may be more cognitively ergonomic (i.e., better affords conceptual metaphors and movements consistent with integer arithmetic) than a chip model. In order to capture the complexity of students' initial learning of integer arithmetic with just one model in ways that addressed limitations noted in prior research, this pre-post-delayed posttest study used multiple methods to address the broad questions: After using either a chip model or a number line model that emphasizes opposites and magnitude as a first integer model, what do students demonstrate knowing about integers and what, if any, differences in learning are found between students who used each model? This article reports the main analyses related to the following specific questions:

1. Prior to instruction, which, if any, conceptual metaphors did students express for integer arithmetic?
2. Which model as a first model better supports overall achievement in the short-term or in the longer-term?
3. Which model as a first model better supports different aspects of integer knowledge (ordering numbers, primary arithmetic operations, generalizing about operations, opposite sums, opposite operations)?
4. Is there evidence that the consistency of model-movements with mathematics influences learning?

Method

I randomly assigned the eight classes of fifth and sixth grade students within two school sites to a chip model or number line model. I did consider including a control group who would only take the assessments to control for repeated testing effects and maturation. This approach was rejected for ethical reasons, such as how this might inequitably prepare control students for next year's instruction.

Settings and People

Research personnel and positionality. Although I position myself as a teacher first and researcher second in terms of wanting students I instruct to learn as much as possible, my role in these districts was as a guest researcher who taught intact classes. Thus, I refer to my role in these classrooms as Researcher-Teacher to foreground the researcher role. As the researcher-teacher, with approximately 20 years experience teaching mathematics (including having taught integer arithmetic to K-16+ students), I conducted all instruction and administered all tests to all students in the targeted grade. The classroom teacher remained in the classroom to ensure safety of students, but not to teach during the study instruction. The study instruction was the only mathematics lessons students at both sites experienced each day. During the five weeks between the post and delayed posttests at both sites the classroom mathematics instruction focused on geometry. During school instruction between these test administrations, students did not use negative numbers, nor were they notified ahead of time that they would be retested, so students did not review integers prior to the delayed posttests.

I have worked for more than two decades as a teacher and researcher using and considering how students use the models compared in this study. I used a chip model when I first taught Grade 8 students integer arithmetic working with a master veteran teacher affiliated with a university in an urban context. The students I taught early in both urban and private school contexts helped me begin to see specific challenges of learning with both typical chip and number line models. In response to challenges of the typical number line model, I develop the walk-it-off model (Nurnberger-Haag, 2007). Although the chip model was the first model I used to teach students integer arithmetic, at the time of this study, I had more experience teaching students using the walk-it-off model.

As I conducted this study, I designed each model to be a “good” representative of the larger class of models of which they were examples (chips models and number line models). A chip model is a commonly used instructional model; many teachers believe in its value. Had I believed the chip model to have little value for student learning, I would not have used it. In any research study, I believe my first ethical duty is to do the best I can for any students with whom I work. I conducted this study in the hope of contributing to knowledge about student learning for all students, but first and foremost my intent is to if at all possible, benefit the students with whom I worked and the teachers who might benefit from observing a guest teacher in their classrooms.

Given the theoretically predicted difficulties with the chip model and the fact that I had more experience teaching with the walk-it-off model, I made additional efforts to ensure that the chip model instruction was as effective as possible (e.g. attended a CMP workshop about effective chip model instruction and sought advice from teacher-leaders who strongly advocate the chip model).

Sample size, settings, and participants. In spite of these commitments, in order to experimentally control for potential perceived or unintentional researcher bias, the study design originally included a second researcher-teacher who was blind to the study theories to teach half of each integer model classes. A power analysis, however, indicated the necessary sample size would be logistically prohibitive to conduct the study as planned. A power analysis for 80% power suggested that if a single researcher-teacher implemented the study instruction, a sample size of 50 in each of two conditions should make it possible to detect between group differences using Cohen’s d 0.25 level (Cohen, 1988).

Two public rural districts in one Midwest state participated after recruiting districts that met the criteria within about a one-hour travel radius of three universities in two Midwest states. To investigate how each model might impact learning if used as the first model students encounter, the study instruction occurred in the grade prior to when integer arithmetic was typically taught in each district (District A first semester sixth grade, District B second semester fifth grade). All district students in this grade attended the same school with the same mathematics teacher in classes that were not ability grouped. The five classes in District A used *Connected Mathematics Project 2 (CMP)* textbook (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006) and the three classes in District B used enVision Math Grade 5 (Scott-Foresman, Addison-Wesley, 2011). According to the state website, 45% of the students I instructed had free or reduced lunch and were primarily European American. The teachers determined how students' effort during study instruction counted toward students' daily participation grades, but individual test scores did not count toward students' grades as it would with typical classroom instruction. Only those students who themselves assented and whose guardians consented participated in the research by giving their written work and assessments to the researcher-teacher for an incentive the equivalent of a university folder and pencil. After removing students due to absences, 164 students remained in the analysis (83 in chip condition and 81 in walk-it-off condition). All possible data were used for each analysis, so the sample size of each analysis varies depending on which students missed particular tests.

Instruction

I planned and implemented a unit with eight lessons of approximately 50-minutes within their normal classroom contexts. I taught students in the ways I would teach if I were their classroom teacher with the following exceptions to avoid cross-contamination between models: I did not assign homework and asked students not to talk about negative number arithmetic outside of the classroom. The first four columns of Table 3 describe the lesson sequence common to both integer models. Each class experienced parallel instruction with the same tasks and activities, differing only by the integer model used (the physical representations, the language about how to use those representations and corresponding model-movements). Subsequent sections detail unit differences due to the integer model, whereas the last column of Table 3 describes only those additional model differences specific to the particular tasks and activities.

Lesson sequence and tasks. I implemented eight days of lessons at Site A, which used 55-minute periods. When I added Site B where class length varied from 30 to 45 minutes depending on the day, I implemented the eight lessons to parallel total time at Site A. The instructional units addressed negative numbers, ordering numbers, and operations with negative integers (addition, subtraction, multiplication, and division) including sums of additive inverses. The unit began by establishing class norms (e.g., what it means to help another student during group work), introducing the integer model representations, and ideas of numbers that were opposite of whole numbers (e.g., opposite color chips with the same cardinality or points on the opposite side of zero with the same distance from zero). Although students discussed and modeled what the opposite of a number meant in words and numerals, they were not introduced to opposite as an operation nor the mathematical notation -(-5) for example, so that these tasks could serve as transfer tasks on the written and interview measures to assess how students would make sense of such notation after using each model. Subsequent lessons involved each pair of inverse operations together, first addition and subtraction then multiplication and division. At the end of each lesson I asked students to complete an exit ticket, but I will not analyze the contents of that data here.

Groupings and approaches. At both sites when seated, students sat at assigned tables. Students were already familiar with working in assigned groups to learn mathematics. During each lesson students worked on tasks and played games in assigned groups and had the opportunity to participate in whole-class discussions. I used students' pretest profiles to assign heterogeneous trios or pairs, yet without large achievement disparities. For most of the unit, the only individual work-time consisted of completing the end of class exit ticket. During the last few lessons, instruction began with independent work and then group discussion until consensus reached or the group agreed they needed to ask for teacher assistance.

I assigned group roles that students rotated after each task to ensure each student repeatedly had the chance to enact the mathematics through oral words, physical movements, and written notation (Nurnberger-Haag, 2007). The group roles were Director (verbally suggested how to use the chips or walk to solve the problem), Chip Handler or Walker (the only person during that task who was supposed to touch the chips or walk the number line to ensure each person had the opportunity to do so and to encourage the need for students to communicate about how to think about and experience the model-movements), and if in a trio, a Verifier (confirmed Director's directions and chip handler/walker's movements). Each student was responsible for writing solutions as complete equations using formal mathematical notation. Near the beginning of instruction students in both conditions were expected to use the tools of their integer models. Later, to help students transition to thinking about the problems in ways that they would be asked to on the tests as well as in future mathematics, I encouraged students not to use the tools and imagine or draw them, if needed.

Chip model lessons. Students used chips on the space in front of them on their desks. With a chip model, numbers are represented by the color and quantity of chips and students can model addition operation as putting in chips and subtraction operation as taking away chips (See column 1 Table 1 meanings). Although I was careful to state that this taking away meaning is just one way of thinking about subtraction (e.g., a difference meaning is also valid), the typical chip model uses this take away meaning of subtraction and is the reason chip models are advocated. Multiplication and division with a chip model uses the idea of repeated groupings in which a number represents the scalar or the number of groups added or removed and the other number represents the quantity in each of those groups.

In a chip model, the only feature of the chips that differentiate different qualities is the color. In order to begin with student thinking using colors, black chips were used to represent the counting numbers with which students were familiar and students explained that white was the opposite of black. Thus, instruction began with white chips representing numbers that were the opposite of numbers most students referred to as "regular numbers." On day three, I offered students the choice to continue using white or begin using red chips to represent negative numbers, because in business contexts red signifies negative quantities.

Walk-it-off model lessons. Just as black and red colors were used in the chips condition, during initial walk-it-off model instruction students recorded positive numbers using a black marker, negative numbers in red, and 0 in some other color. Each group used wet-erase markers to write on a ten-foot long open number line (line permanently drawn on white plastic tablecloth strips). The students in the chip condition were able to begin addition and subtraction earlier in the unit, because they more easily associated the sets of opposite color chips with opposite numbers than the students who had to learn how to and spend time to construct number lines. To ensure similar experiences between classes and groups, all students placed the number lines on the floor and I asked them to walk next to the number line.

Different ways students moved on the number line to represent the expressions were discussed as a class. Addition and subtraction problems in which the first number is a starting point or location just as it is in whole number arithmetic. Just as with a chip model particular meanings of operations are used such as take-away or add in, the walk-it-off model works for all operations because it uses an opposite direction meaning of the "-" symbol, so when students moved in the opposite direction on initial problems, but referred to it as going backwards or to the negatives etc. I encouraged students to think about the meaning of "-" on a number line as meaning turn the opposite direction.

Data Sources

The primary data sources to answer the research questions were collected via two written open-response measures a skill-based *Integer Arithmetic Test (IAT)* and *Explain & Draw* Test (*EDT*). Students took these tests three times on parallel forms as pretests, posttests, and delayed-posttests. Primary measures were repeated at each testing phase, whereas covariate measures (*Timed Fact Test* and *Spatial Test*) were given at pretest to control for abilities other than negative number arithmetic prior to instruction. Written pretests were administered several days to a week before the integer lessons began. Posttests were given the day after instruction. Absent students were given the tests as soon as possible thereafter. Approximately five weeks after instruction ended, I administered the delayed-posttests at each site with no make-up tests given.

All students took the IAT before the EDT at each test administration in order to ask the skill-based questions prior to the items that asked students to explain, because the act of explaining could change performance on the skill test. To limit the likelihood of student difficulties with reading impacting their performance on these tests, before students began the IAT I read each set of directions to the class pointing to the text on a document camera. After most students were ready to begin the EDT, I read each problem and encouraged anyone wishing to have a problem reread to them to ask me to do so. Details about each test follow in subsequent sections.

Integer Arithmetic Test (IAT). The *Integers Arithmetic Test* assesses integer understanding (e.g., order of integers) and calculation skills (addition, subtraction, multiplication, and division) as is typical of standardized and classroom assessments (e.g., *TIMSS*) and some prior research (e.g., Liebeck, 1990; Periasamy & Zaman, 2009). The exact problems used during instruction were not replicated on any test in order to ensure that students did not simply remember how to solve a particular task (with the exception of the open-ended task in which students added and subtracted to make 0). Although more complex problems were considered during the test development, all operations consisted of only binary operations. This was done in order to create a set of tests that could fit within the practical time constraints of a class period yet include a full range of binary problem structures, so additional types of problems were left for future research. The IAT consisted of five ordering number items, 10 each of addition and subtraction calculation (and two addition of the form _ + _ = 0, *Generate Opposite Sums* items), eight multiplication and seven division (because one item was removed due to not working properly), five opposite operation problems (three opposite of a negative number and two opposite of an expression).

The written measures assessed the constructs taught during instruction as well as opposite operators (transfer problems) that were not taught during the lessons (e.g., -(-4) and – (6-8)). The data reported here are from a 46-item open response skill-based test *Integer Arithmetic Test* (IAT) and a seven-item *Explain and Draw Test* (EDT), three items of which are included in this analysis. These measures were developed through several phases of piloting and analysis including factor analysis to remove items that did not perform as expected. Due to the three test administrations, I developed three equivalent forms, which an ANOVA confirmed were not significantly different. Although no significant differences were found between forms, to experimentally counteract potential order effects of the three administrations, the students were randomly assigned to one of six possible pre-post-delayed posttest form orders prior to the pretest. Cronbach's alpha is an appropriate measure for internal consistency for unidimensional items (Tavakol & Dennick, 2011). For each unidimensional subconstruct, Cronbach's alpha was greater than .7 (.74 to .98), with the exception of two subconstructs. One subconstruct consisting of four items composed of two subtraction structures empirically loaded onto one factor and fit theory based on the problem structure but this loading was small and loaded onto two additional factors as well (4 items, α=.59) The other subconstruct only consisted of two items, which were the two most difficult transfer opposite operation tasks (.42). Pencils were the only tools allowed during the written assessments. I decided not to provide chips or number lines, so that the tests were able to measure students' use of formal mathematical notation and reasoning as will be tested with most classroom and standardized assessments.

Explain & Draw Test (EDT). The seven-item *Explain & Draw test (EDT)* was designed to obtain insights about students' reasoning, including their use of metaphor and movement in their conceptions of integer arithmetic. Although this type of measure may not be as effective as an interview for this purpose, the format of a written test allowed for larger sample sizes in order to assess what might be typical of students in each model condition. The EDT asked students to explain in words and by drawing how they make sense of the ideas in each question so that another student their age would understand. See Table 4 for a summary of the purpose of each question with examples. The items were designed to investigate if students would offer accurate generalizations about integer arithmetic. The *Operation Generalization Item*, which was revised from the 2003 eighth-grade *TIMSS* released item discussed in the review of literature to be more appropriate for fifth to seventh grade students, assessed whether students continued to generalize based on whole number arithmetic such as subtraction always makes numbers smaller.

Covariate measures. To control for calculation speed and obtain a proxy for mathematics achievement prior to instruction without accessing students' personal records, students took a *Fact test*. I constructed this 1-minute whole-number fact test of the four primary operations using digits from the items on the three forms of the IAT. Students also took two spatial tests (Ekstrom, French, Harman, & Dermen, 1976); however, these tests are not analyzed here and future research could consider whether spatial ability was a mediator of learning for each model.

Pilot test results revealed an additional covariate would be needed when assessing multiplication and division. If students consistently answered all multiplication and division pretest items with negative or with positive solutions, this falsely inflated their pretest score such that real learning gains at posttest and delayed posttest could be underestimated. Consequently, a categorical covariate variable of MD Preconceptions was added, using 1 for presence and 0 for absence of these particular preconceptions. I documented presence of MD Preconceptions if students answered at least 14 of the 15 multiplication and division items either consistently positive or consistently negative. The significance of the MD Preconceptions in the current study analyses confirmed the need to control for this factor in the analyses.

Data Analysis

Statistical controls were built in to the study design and analysis (e.g., Hill & Shih, 2009), including a pretest—the best covariate in educational research (Shadish, 2012), fact test, gender, district, and preconceptions of negative number multiplication and division. Although district and class were both originally included as statistical controls, the statistical models would run with only one of these and either covariate resulted in the statistical model accounting for the same amount of variance. Student grade level also differed between districts so differences due to grade cannot be separated from district differences. Multivariate analyses of variance on the two dependent pretest measures of Pretest IAT and Fact Test found statistically significant differences between districts ($p = .010$), but no significant differences were detected between the eight classes or between assignment of integer models prior to instruction. Thus, I included district in each analysis that used covariates. The analyses were robust to violations of assumptions, because the integer model group sizes were similar. Unless otherwise explained in the results, data for which ANCOVA, MANCOVA, ordinal or logistic regressions were used satisfied the assumptions of the stated test. Non-parametric tests were used when assumptions were violated (ordering numbers analysis). Interrater agreement for each qualitative analysis was greater than 90%.

Total IAT and primary operations analysis. The IAT total test score was calculated using 0 for incorrect items and 1 for correct then scaled to a 100-point test to weight the relative importance using the subtotals of the following constructs: ordering numbers (20%), addition and subtraction (35%), multiplication and division (35%), and opposite operations (10%). Integer model group performance on posttests and then delayed posttests were separately compared using an analysis of covariance, controlling for the factors described in the overview.

Raw scores of the primary operations were then compared to test for posttest differences between consistent and inconsistent model-movement problems due to integer model using a MANCOVA, controlling for these scores at pretest, fact test, gender, district, and MD preconceptions. Delayed posttest consistent and inconsistent model-movement raw scores were then similarly analyzed. Degree of model-movement consistency was also assessed using categories of model-movements found with the models: Inconsistent Beginning Model-Movements, Inconsistent In-Process Model-Movements, and Consistent Model-Movements. Each of these categories had different total possible scores (10, 4, and 22 respectively). Given these unequal scales, I standardized the values in order to be able to compare parameter estimates to test if the greater the degree of consistency the greater the accuracy. Then a MANCOVA was conducted on the standardized posttest scores controlling for these scores at pretest, fact test, gender, district, and MD preconceptions. Standardized delayed posttest scores were then similarly analyzed.

Opposite sums knowledge. Multiple methods were used to determine each student's level of opposite sums knowledge at pre, post, and delayed posttest using three types of tasks: IAT calculation item accuracy (e.g., single digit -7 + 7= and double digit -19 + 19=), IAT generative item accuracy (___ + ___ = 0), as well as accuracy and reasoning on the EDT Opposite Sums Item. I coded student explanations of the EDT Opposite Sums item with a qualitative coding scheme developed from the pilot data using a constant comparative approach (Glaser, 1965). The codes used in the analysis of the full study data reported here documented whether students' explanations provided evidence that they had generalized that the sums of any opposite numbers was zero, whether they calculated to explain their answer, stated that the quantities could not be equal because different numbers always create different solutions, no reasoning, and other less frequent codes. A second trained coder assessed 20% of the randomly selected codes with 92.7% agreement. These codes were developed through at least three phases. First, I developed an initial coding scheme based on test-only pilot data from seventh grade students who had studied negative number arithmetic in full-length units with their own teachers using both chip models and number line models. Second, I tested and revised the coding scheme using the pre-post pilot study I conducted with fifth grade students. Final revisions occurred with the full study data.

Each student's EDT Opposite Sum reasoning code was then combined with the accuracy score into an EDT Opposite Sum item level (0-3). Two types of IAT questions (two Calculate Opposite Sums items such as -19 + 19 and two Generate Opposite Sums items ___ + ____ = 0), were each separately subtotaled for accuracy 0 to 2. I used a K-cluster analysis procedure using these three ways of demonstrating opposite sum knowledge (calculation subscale 0-2, generative problems subscale 0-2, and EDT Opposite Sums Item Level 0-3) to inform selection of Leveled Opposite Sums knowledge profiles (no/low, moderate, or strong). Figure 1 illustrates the sequence of steps to mix these data sets for statistical analysis. The post levels were then analyzed using ordinal regression to compare integer model differences and a separate analysis conducted for delay post levels controlling for Opposite Sum Knowledge Level at pretest and the other controls included in the total test analyses.

Ordering numbers. The accuracy of the five IAT items that involved ordering integers were combined to create an Ordering Numbers score. These items consisted of three types: Circle the largest number, circle the smallest number, and give two numbers that are smaller than N. In part, due to ceiling and floor effects creating bimodal distributions, the distributions did not meet assumptions to conduct a MANCOVA to compare the integer models. Transformations did not correct these violations. Consequently, non-parametric tests (Mann Whitney) were conducted separately on posttest and delayed-post ordering scores to compare integer models.

EDT generalizing operations item. Accuracy on the posttest and then the delayed posttest EDT Generalizing Operations item were analyzed using logistic regression. The student explanations on this item were coded with a qualitative coding scheme developed from the pilot data using a constant comparative approach (Glaser, 1965) in light of theories of symbol use (Vlassis, 2008) and generalizing of operations. Related codes were combined to yield three primary categories of reasoning: (a) Accurate Generalized rules or meanings, (b) Calculated, or (c) Overgeneralized from whole number operations. A Chi-Square Test of Independence for integer model was then conducted for type of reasoning.

Opposite operations. The three types of tasks from the IAT and EDT that assessed meanings of "-" as opposite operations were used to determine each student's pre, post, and delayed post Opposite Operation Knowledge Level. The accuracy on the two types of tasks on the IAT (*Opposite-Single Digit* and *Opposite Expression*) that asked students to calculate with specific numbers were leveled as low (0) or high (1) performance. Student reasoning on the EDT Opposite Operation Item that asked students to generalize about the opposite of a negative number was qualitatively coded to reveal types of reasoning. These types of reasoning were then grouped and assigned a level 0 to 3. Rather than deciding apriori how students' should be discriminated by performance across these three task types *Opposite-Single Digit, Opposite Expression*, and *EDT Opposite Operation Item*, I used a K-Cluster Analysis procedure on posttest scores to inform levels of opposite operation knowledge and identify profiles of student knowledge within each level. I intended to analyze these levels with ordinal regression, controlling for several factors; however, the sample size was not sufficient for enough for students to fall within all the multiple classifications determined by the factors to perform a reliable analysis.

Conceptual Metaphor Analysis

Conceptual metaphor analyses were used to determine which, if any, conceptual metaphor students expressed prior to integer arithmetic instruction. Thus, a coder blind to the study hypotheses analyzed EDT pretests, and I report these descriptive statistics. Intervals of the randomly assigned student identification numbers were selected for this detailed analysis. Since the descriptive results of this subset of students were consistent with analyses using this coding scheme on EDT tests with another sample (Nurnberger-Haag, 2013), this subset of 38% of student tests was sufficient. I report descriptives of the three items on which students most often expressed a conceptual metaphor (Ordering Numbers item and single digit subtraction and multiplication operation). The coding scheme analyzed written language and drawings of the EDT, which was developed prior to and refined on the test-only pilot data with seventh grade students (Nurnberger-Haag, 2013). Appendix B summarizes these coding definitions for the following behaviors: *Collecting Objects*-Things (if referred to things, objects, or groups), *Moving on a Path*-Points (if referred to numbers as points or locations on a path) or *Moving on a Path*-Path (including distances moved on a path), *measuring* if students referred to static distances as intervals. Coders documented 1 or 0 whether each test item provided evidence of each behavior. I interpreted that a student expressed a particular metaphor if the coder noted a related behavior in words or drawings on any item. I coded 20% of the pretests blind to student condition due to random assignment of the identification numbers with greater than 90% agreement.

Results and Analysis

Changes in test outcomes from pre to posttest and pre to delayed posttest form the bases of these analyses. First I report the IAT test total score using quantitative analysis, from the open-answer skill-based questions of all constructs. Next each construct was analyzed using multiple methods to explore student expression of construct knowledge based on accuracy of each type of construct question on the IAT as well as accuracy and reasoning on the related item of the EDT. These constructs include overall primary operations (addition, subtraction, multiplication, and division), sums of additive inverses (a particular case of addition), ordering numbers, and opposite operations.

Lastly, to investigate whether evidence supports prior claims that a chip model would be more likely to draw on students' intuitive prior knowledge of whole numbers in beneficial ways, additional analysis of EDT responses prior to instruction were analyzed for evidence that students would naturally express ideas of negative numbers as collecting objects as resources for thinking about negative number arithmetic.

IAT Total Score Analysis

Changes in performance within model. Figure 2 displays the pre, post, and delayed posttest unadjusted means with error bars that show the 95% confidence intervals. Note that the unadjusted pretest means of each integer model were similar (chip M=39.0, SD=13.3 and walk-it-off M=41.2 SD=16.5). The unadjusted means for average achievement of students in both the walk-it-off model (post M=67%, SD=19%; delayed post M= 64%, SD=23%) and the chip model (post M=53%, SD=20%; delayed M=49%, SD=20), as well as Figure 2, show that on average students who used both model improved. These improvements were statistically significant ($p<.001$), indicating that after eight lessons both models contributed to student learning. These scores include 10% weighting of a transfer construct on which most students were not successful and would not normally be factored into a classroom grade, which skewed the test scores down compared to classroom assessments that typically test only instructed material. Thus, a 90% on this assessment might be considered a perfect score on instructed material.

Between integer model differences. Analyses of covariance (ANCOVA) were conducted on the IAT post and also on delayed posttest total scores to compare students' overall learning with the walk-it-off model to learning with a chip model controlling for pretest IAT, whole number fact test, gender, district, and preconceptions of integer multiplication and division. Table 5 displays the model statistics and Table 6 provides the parameter estimates of this overall analysis. As shown in Table 5, the statistical models of each ANCOVA including all significant (Integer Model, Pretest IAT, Gender, District, MD Preconceptions) and a non-significant predictor (Fact Test) accounted for approximately 46% of the variance (Adj. R^2=.458) $F(6, 152)=23.24$, $p<.001$ on the posttest and 47% (Adj. R^2=.468) $F(6, 149)=19.847$, $p<.001$ on delayed posttests. The next greatest predictor after the pretest was the independent variable under investigation, the Integer Model. Statistically significant differences on posttests were found between students who used the walk-it-off compared to chip model: $F(1, 152)=17.110$, $p<.001$, $\eta\rho^2=.10$) and also on the delayed posttests $F(1, 149)=18.8$, $p<.001$, $\eta\rho^2=.11$.

As the parameter estimates indicate in Table 6, on average on this 100-point test students who learned with the walk-it-off integer model scored 10 points higher than students who had learned with chips on the posttest and 12 points higher in the longer-term on the delayed posttest. Just as in classrooms some types of problems are deemed worth more points than others, the IAT test constructs were weighted as described in Methods. Consequently, although these results cannot be directly interpreted as number of questions correct (and consequently, students with the same score could have different numbers correct), these scores in Table 6 could be interpreted in terms of achievement as percent grades.

Integer Operations

In this section, I first report the accuracy and reasoning results of the *Generalization Item* of the EDT, a multiple choice item that also asked students to explain and draw their reasoning. It required students to consider how all four basic operations work (revisit Table 3). Next the basic operation calculation items of the IAT were further analyzed. I first use the theoretical perspective of model-movement consistency to group these operation items. I then discuss performance on the individual operations (addition, subtraction, multiplication, and division) in relation to this analysis.

Generalizing about operations.

Generalization item accuracy. Five-weeks after instruction, 32.1% of 81 students who learned with the walk-it-off model compared to 10.0% of 80 students who learned with the chip model correctly answered the EDT *Generalization Item.* I conducted logistic regression on the accuracy of the *Generalization EDT Item* to test for integer model differences controlling for the Pretest Generalization Item Accuracy, gender, district, and fact test. When controlling for these predictors, using the walk-it-off model significantly increased the odds of accurately answering this generalization question by a factor of 6 (Exp(B)=5.98, $p<0.001$). No significant differences on generalization accuracy were found, however, for integer model at posttest (30.5% walk, 22.6% chip).

Generalization item reasoning. The reasoning students offered differed significantly on both post and delayed tests. Considering only the reasoning, students who learned with chips (a) used a Calculation Strategy (irrespective of accuracy) 2.5 times more often than students who learned with the walk-it-off model, meaning students substituted a negative number into each of the four expressions to calculate answers to determine an answer choice and (b) 1.6 times more often used an Overgeneralization Strategy from whole number operations. In contrast 9 times more frequently students who learned with the walk-it-off model offered an accurate generalized explanation tied to the meaning of operations with negative numbers or a rule students had generated (not explicitly taught during instruction).

As described in methods, students primarily reasoned about the generalizing operations item in the following ways: Calculated Expressions (regardless of accuracy, entered a single number into each answer option and calculated to determine the greatest number), Overgeneralized from whole number operations (e.g. "because multipling makes the number bigger" student 3092, spelling as student did), or provided Accurate Generalized Rules or Meanings. Some accurate generalized rules or meaning examples include: "In negative #s you subtract the more you have" ID 3170. "Subtracting would give you a larger number because you take away the negative number and get closer to zero or a positive number" ID 3164 "For subtraction you have to turn the opposite direction twice if you have a negative number so that means you're walking the positive direction." ID 3130).

A Chi-square test of independence compared the qualitative reasoning of the 74 students who offered codable reasoning on the posttest by integer model. Fisher's Exact Test (used due to one cell being less than 5) found a significant interaction of reasoning and integer model on the posttest $p<0.001$. In contrast to the posttest, on the delayed posttest, students who learned with either integer model were equally likely to use the Calculate Expressions Strategy described above (chips 19/38; walk 20/41). Students assigned to walk-it-off, however, still used accurate generalized reasoning (9/41) 9 times more frequently than students assigned to chips (1/38) and were less likely to over generalize from whole number operations with statements like "multiplication always gives the biggest answer" (chips 18/38, walk 12/41 27%). A significant interaction of reasoning and integer model were found with a Chi-Square Test of Independence on the delayed posttest ($p=.021$, using Fischer's Exact Test due to one cell being smaller than five).

Consistency of model-movements with integer operations. Although analyzing test items grouped by operation is a common educational approach, I grouped problems by the physical model-movements to offer new insights for analyzing model-efficacy. As discussed in the theoretical framework and methods, some IAT problems require students to move chips in ways that are inconsistent with the targeted integer operations.

Consistent vs. inconsistent model-movements. Multivariate analysis of covariance (MANCOVA) was conducted on the IAT operation posttest Consistent Model-Movement Problems and Inconsistent Model-Movement Problems controlling for these scores at pretest as well as the other controls used in the total IAT score analysis. The corrected statistical models for the consistent item posttest scores were significant $F(7, 152)= 9.169$, $p<.001$, Adj. $R^2=.265$ and inconsistent $F(7, 152)=17.303$, $p<.001$ indicating this set of factors explained variations in student scores. Multivariate tests found integer model as well as all covariates except district were statistically significant at $p<.05$. Table 7 reports the multivariate values. Statistically significant differences approaching a large effect size (Cohen, 1988; Tabachnick & Fidell, 2013, p.87) were found on those posttest problems for which I argue that the chip model required inconsistent movements—$F(1, 152)=42.861$, $p<0.001$, ($\eta\rho^2= .220$). Multivariate analysis of covariance (MANCOVA) on delayed posttests using the same covariates were also statistically significant approaching a large effect size: $F(1, 148)=46.496$, $p<0.001$, ($\eta\rho^2= .239$).

Table 8 provides the parameter estimates of this analysis. Each point represents one problem, so the parameter estimates (see β and 95% CI in Table 8) indicate that compared to students who used the walk-it-off model, out of 14 problems students using the chip model accurately answered about 2 to 5 fewer problems (posttest $\beta=-3.24$; delayed posttest -3.54). When the chip model-movements were consistent with the mathematics of integer arithmetic operations, however, no significant differences were found between the integer model groups' test performances. Note integer model groups differed by less than one out of 22 Consistent model-movement problems (posttest: β=.-.843, delayed posttest (β=-.897).

Degree of model-movement consistency. Multivariate analysis of covariance (MANCOVA) on standardized scores of posttest IAT operation Inconsistent Beginning Model-Movement problems, Inconsistent In-Process Model-Movement Problems, and Consistent Model-Movement Problems was conducted to test if the degree of model-movement consistency mattered controlling for these scores at pretest as well as the other controls used in the total IAT score analysis. See Table 9 for the multivariate test values and Tables 10 and 11, respectively for parameter estimates of posttest and delayed posttest degree of model-movement consistency. The posttest differences were statistically significant with a large effect size of integer model (chips n= 82, walk = 77, see Table 9). Between subject effects were significant for Inconsistent Beginning $F(1, 150)=45.92$, $p<0.001$ ($\eta\rho^2= .23$) and Inconsistent In-Process problems $F(1, 150)=7.519$, $p<0.007$ ($\eta\rho^2= .05$), and not significant for Consistent problems. MANCOVA on standardized delayed posttest scores of the same problems were also statistically significant with a large effect size of integer model (chips n= 78, walk = 77 (see Table 9). Between subject effects were significant for Inconsistent Beginning $F(1, 146)=51.139$, $p<0.001$ ($\eta\rho^2= .27$) and Inconsistent In-Process problems $F(1, 146)=7.118$, $p<0.008$ ($\eta\rho^2= .05$), and not significant for Consistent problems. These results may reflect a large effect size for Inconsistent Beginning and small effect size for Inconsistent In-Process. The more closely the chip model-movements aligned with the mathematics, the more likely students who used chips were to be successful on those problems types, leading to less significant differences when compared with similar peers who used the walk-it-off model. Standardized scores were used to create outcome variables with the same scales in order to compare parameter estimates for effect of consistency. Note in Tables 10 and 11 that both for posttests and delayed posttests the parameter estimate disparities for integer model performance were greater the more inconsistent chip model-movements were with mathematics.

Sums of Additive Inverses (Opposite Sum) Knowledge

Evidence of students' opposite sums knowledge expressed on the EDT Opposite Sum item, two IAT Calculate Opposite Sums items and two IAT Generate Opposite Sums items were combined as described in the Methods section to classify each student's level of opposite sums knowledge (1, 2, or 3). Ordinal regression analysis on the 160 students' levels of knowledge for whom a level could be determined (i.e., data was not missing), showed that using chips did significantly increase the chances of demonstrating greater opposite sums knowledge at posttest than the walk-it-off model (p=0.001, pretest opposite sum knowledge p<.001, fact test p=.002, district n.s. and gender n.s.), but this difference was not maintained five weeks later. There were no significant differences between students who learned with the chip model or walk-it-off model on the delayed posttest (n=156, p=0.096).

Ordering Integers

An analysis of pre to post and pre to delayed post changes in Ordering Integers subscores revealed that twice as many students who learned with chips regressed compared to students who learned to walk-it-off on both the posttest and delayed posttest. Table 12 presents descriptive statistics on the frequency with which students improved, maintained, or regressed on the five ordering items from pre to post and pre to delayed posttest. As described in methods, the post and delayed posttest scores were not normally distributed, so Figure 3 displays box plots of the pre, post, and delayed posttest ordering numbers totals (0 to 5 possible), which provide a more appropriate representation than means and standard deviations. Figure 3 shows that a majority of students from both models were at or near ceiling at pretest. As the median of the box plot indicates, at posttest and delayed posttest at least 50% of students in both integer model conditions answered all five ordering problems accurately. Note that the box plots also reveal that the lowest quartile of students who learned with chips correctly answered 0 to 3 problems at post and delayed, whereas students who learned with the walk-it-off model that had these low scores were outliers. A Mann Whitney Test found these differences between integer models on the total posttest ordering scores not significant, but students who learned with the walk-it-off model did significantly better than those who learned with chips on the delayed posttest ($p=.018$).

Opposite Operation

Ordinal regression analysis was not reliable on the transfer construct level of Opposite Operation knowledge, because only 12 out of 162 students were successful on these items. Thus, as described in the Methods section, statistical analyses for the opposite operation construct are not reported.

Students' Initial Metaphors

All results thus far have been post and delayed post achievement controlling for pretest scores and other variables. In this section, I first use results only from the pretests of the EDT to provide evidence of students' reasoning prior to instruction. Because the EDT test items did not prompt students to use a conceptual metaphor, there was no reason for students to express ideas consistent with a particular conceptual metaphor unless this was their current way of thinking. To find empirical evidence that students find thinking of numbers as objects (Küchemann, 1981; Liebeck, 1990) or as moving on a number line more intuitive (Freudenthal, 1973), I analyzed a subset of the pretest EDT data (as described in Methods). Consequently, this analysis does not compare students by condition, but investigated what 63 (38%) students thought prior to instruction. As Table 13 shows for each type of assessed problem, students most frequently used the moving-on-a-path metaphor. Moreover, a majority of the students who expressed any metaphor used the moving-on-a-path metaphor. The collecting objects metaphor was next most popular and measuring metaphor least popular.

Do Students Need Integer Knowledge to Benefit from Chip Model?

Particularly if prior to instruction students do not typically begin thinking about negative numbers as collecting objects, and because the chip model requires students to accept and use the idea that positive and negative objects cancel for most problems, I predicted that some level of opposite sum knowledge might be necessary prior to instruction in order to learn with a chip model. To test if those students who prior to instruction did not have strong sums of additive inverses knowledge had more difficulty learning with this chip model than the walk-it-off model, an analysis of covariance (ANCOVA) was conducted on the posttests of participants (n = 122, chips 65, walk 57) whose opposite sums knowledge was not strong at pretest (level 1 or 2). The overall model was statistically significant $F(6, 115)= 13.452$, $p<.001$ and accounted for about 38% of the variance (Adj. $R^2=.382$). Such students assigned to the chip model did significantly worse overall on the IAT than students assigned to the walk-it-off model: $F(1, 115)=17.984$, $p<0.001$; $\eta\rho^2=.135$. Similarly, on the delayed posttests (n=120, chips 62, walk 58) the statistical model was significant $F(6, 113)=14.243$ $p<.001$and accounted for 40% of the variance (Adj. $R^2=.382$). Students who used chips did worse overall than those who used the walk-it-off model $F(1, 113)=23.976$, $p<0.001$; $\eta\rho^2=.175$. These students with lower opposite sum knowledge who used chips scored lower on the 100 point posttest -12.5, 95% CI [-18.3, -6.6] and delayed posttest -15.4, 95% CI [-21.7, -9.2].

Discussion

Both models supported statistically significant student integer learning from pre to post and pre to delayed posttest, indicating that both models are reasonable instructional models to use for integer instruction. As the first model students encounter, however, the walk-it-off model helped students outperform students who used the chip model on short-term as well as longer-term skill tests that assessed instructed aspects of integer knowledge (ordering numbers, all four primary operations, and sums of additive inverses,) and transfer knowledge (opposite operations). These statistically and practically significant results indicate that relative to this chip model, the walk-it-off model was a more effective initial model to support students' first-time integer arithmetic instruction. The delayed post-test results were especially noteworthy given that the classroom teachers' subsequent lessons did not use negative numbers, and students did not review integers prior to the delayed posttest. Analyses of each aspect of integer knowledge that mixed qualitative assessment of student reasoning with skill-based performance provided additional evidence for these conclusions and a more nuanced understanding of what each integer model might better support. Table 14 summarizes which, if either, integer model was more advantageous for each aspect of integer knowledge at each time point.

Overall the walk-it-off model was at least as good, if not a better first instructional model for each aspect of integer knowledge investigated (see Table 14). For example, although in the short term students working with the chip model did outperform their peers who learned with the walk-it-off model with regard to sums of additive inverses knowledge, the walk-it-off model was equally effective in the longer-term. With regard to ordering numbers, because a number line model orders numbers, we might expect that students who used walk-it-off would have outperformed students who learned with chips. Walk-it-off was more advantageous, but only in the longer term. This difference widened because twice as many students who learned with chips than the number line regressed on ordering knowledge from pretest to delayed posttest.

The generalizing operations results demonstrate potentially even more powerful walk-it-off model impact on student performance than for primary operations. On the generalization item revised from the TIMMS assessment with an international average of 40%, delayed posttest accuracy of these younger students (fifth and sixth grade) who used the walk-it-off model (32%) were closer to the eighth grade international average than students who used chips (10%). In other words, the walk-it-off model not only facilitated better procedural accuracy to first learn primary operations but also led students to more accurately generalize about operations than those students who first learned with chips. If one takes the perspective of algebra as generalized arithmetic (Usiskin, 1988), a purpose of experience with negative number and rational number arithmetic is to help students accurately generalize how operations interact with all real numbers. When these targeted generalizations contradict their experiences with whole number operations, students have difficulty accommodating to fit the ways that negative number and rational number arithmetic works (Ryan & Williams, 2007). As Table 14 summarized, in the longer-term the walk-it-off model better supported ideas of generalization both in terms of accuracy and reasoning. This longer-term assessment is more reflective of how students' integer arithmetic understanding might come to bear when studying formal algebra during the next few years. For example, students who learned with the walk-it-off model were less likely than students who used chips to still say that multiplication will always make a number larger or subtraction always smaller.

Some reasons that opposite operation knowledge could not be analyzed may be because few students from either model accurately answered these unfamiliar tasks. This aspect of knowledge warrants additional research with qualitative methods (see Article Two). It also warrants future studies that include this aspect of knowledge in the instructional unit for several reasons. First, grounding an opposite operation for the written symbol "-" by moving the opposite direction is the experiential basis for learning integer arithmetic with the walk-it-off model. Since walk-it-off model did advantage students on arithmetic, and translation of spatial understanding to writing or oral language-based symbols usually follows later, even the delayed test of 5 weeks may have not been sufficient time for students' spatial understanding to translate into language. Second, the trend that five weeks later other short-term non-significant results changed to a walk model advantage. Third, uncovering true impacts requires investigation even years later (Yelon, Ford, & Golden, 2013). Finally, but perhaps most importantly, other research signaled potential delayed benefits of whole-body learning (Hadzigeorgiou, Anastasiou, Konsolas, Prevezanou, 2009).

Understanding the Performance Difference

The findings that the walk-it-off model better supported initial student learning align with reasons drawn from cognitive science, analyses of mathematics, and practical classroom experience. Specifically, the following reasons that students may have found the walk-it-off model a more cognitively ergonomic first integer model will be discussed in more detail: (a) integer model-movement consistency with mathematics matters for learning, (b) model consistency with the conceptual metaphor it promotes, and (c) students find moving-on-a-path a more intuitive conceptual metaphor for negative numbers.

Consistency of model-movements with integer operations matters. Some theoretical work referred to the mathematical alignment of integer models as breaking or requiring model-rules that differ from the mathematics (Star & Nurnberger-Haag, 2011; Vig, Murray, & Star, 2014). This study offers empirical evidence supporting these theoretical arguments that incongruent mathematical alignment does impact students' learning outcomes during initial learning. Moreover, this study offers reasons related to human cognition of mathematics that these breaks are likely problematic. Students moved differently in order to interact with representations to enact these models. If these differences and the ways model-movements represent mathematics were not relevant for first-time student learning, then there should be not have been significant differences in performance when problems were grouped by model-movement consistency with integer operations. The analysis of student performance on the primary operations, however, did show that when the chip model required model-movements that contradicted or were extraneous to the integer arithmetic processes and ideas (e.g., adding in chips to subtract) this interfered with learning. In contrast, when a student could engage in integer arithmetic problems using model-movements consistent with the mathematics, either model was equally effective in supporting student outcomes. This lack of significant differences when both models encouraged student movements consistent with the mathematics supports the claim that the model-movement alignment with mathematics is a likely factor of student learning with models.

Perhaps more important for theory and for practical instructional decisions, beyond this dichotomous issue of consistent versus inconsistent model-movements, the findings demonstrated that the greater the degree of model-movement's inconsistency with primary integer operations, the greater the differences found between integer model performance. This finding about integer arithmetic complements other findings about consistency of students' gestures with mathematical symbols (Goldin-Meadow, Cook, Mitchell, 2009). In other words, students found greater model-movement consistency more cognitively ergonomic from a performance standpoint. Future research could ethnographically investigate students' perceptions of these integer models for how cognitively ergonomic they felt.

This issue of model-movement inconsistency in the chip model could be an additional reason that students who learned with chips more frequently regressed on ordering knowledge, beyond the fact that a collecting objects metaphor emphasizes cardinality rather than ordering numbers. When students have to add in chips to perform non-addition operations, their chip model-movements as well as the resulting visual images, change perceptions of the cardinality of the physical objects in ways that would muddle the relative ordering of number values.

Consistency of models with conceptual metaphors matter. Yet, another potential reason that more students assigned to use chips as the first integer model did worse on ordering problems and overall performance on the delayed test could be related to the inconsistency of the mathematics with the conceptual metaphor. Students who used chips and who did not already have strong sums of additive inverses knowledge at pretest scored about 1.5 letter grades lower on the delayed posttest than students who used the walk-it-off model. This may be due to the way in which chip models violate the collecting objects metaphor.

Chip models violate zero-as-nothing. As Rotman (1993) explained, when thinking of numbers using an object collection metaphor, zero should be visually represented as nothing or "no thing." The chip model violates this meaning by requiring multiple things to be "no thing." This zero-as-nothing violation, it should be emphasized, is not limited to the situations in which a mathematical problem uses the number zero. Almost every number that instructional practices ask students to represent with a chip model violates the meaning of that quantity in the collecting objects metaphor. Indeed students need to overcome this zero-as-nothing violation when calculating most problems with chips. For example, to begin the problem -7 – 5 students have to use at least 17 total objects to represent the value -7 (i.e., seven negatives as well as five additional negatives and five positives). It was clear in the pilot study that such zero-as-nothing violations did not feel cognitively ergonomic. For example, one student explained why using -7 and 7 was "kind of zero," but not really: "Right now, there's 14 physical chips, and your brain wants to say they're still there." In contrast, the walk-it-off model does not violate the meaning of numbers represented with the moving-on-a-path metaphor. The findings of this study suggest that because this model requires students to overcome the zero-as-nothing violation to calculate almost every integer problem, opposite sums knowledge may actually be required in order to learn integer arithmetic with the chip model. Thus a chip model should not be the first model that students encounter.

Moving on a path: More intuitive for negative numbers. Research suggests instruction should begin from student thinking (Freudenthal, 1973; Carpenter, et al., 1998). Freudenthal (1973) although advocating more inductive reasoning approaches for integer arithmetic also noted that if a representation were used, a number line would likely be best for students. Yet, some have claimed that a chip model should be a more intuitive way to build on student thinking, because it encourages students to recapitulate the same approach through which they first began to develop whole number understanding (Küchemann, 1981; Liebeck, 1990). With regard to conceptual metaphor theory, this means advocating a collecting objects metaphor. Yet, the pre-post and pre-delayed post as well as prior to instruction results contradict such claims. The fact that most students who expressed some conceptual metaphor at pretest used a moving-on-a-path metaphor suggests it is the moving-on-a-path metaphor that is likely the metaphor students would find more intuitive to begin their work with integer arithmetic. One might argue that the pre-instruction conceptual metaphor results indicate these particular students might be unique and were already using a moving-on-a-path metaphor, such that the students in this sample found the study instruction more intuitive. Future research is needed to confirm that other or most students also find the moving-on-a-path metaphor more intuitive than a collecting objects metaphor for integer arithmetic. Such a theory does seem plausible in light of prior research on student performance after students learned integer arithmetic with their regular classroom teacher using both chip and number line models with CMP in which a chip model was emphasized first (Nurnberger-Haag, 2013). Even when the chip model was emphasized during multi-model instruction, analysis of student performance on the same test items used in this study found more students expressed the moving-on-a-path metaphor than collecting objects post-instruction (Nurnberger-Haag, 2013).

Conclusions

This study suggests that when integer arithmetic with negative numbers is first introduced, the walk-it-off model is preferable to this chip model. As discussed in the theoretical framework, an important feature that distinguishes the walk-it-off model used in this study from other number line models that also draw on a moving-on-a-path metaphor is students learn that instead of signaling a particular direction in space, "-" or "+" signs signal whether to move in the opposite direction relative to one's current direction. Although different chips models are also possible, this study demonstrated stronger performance with walk-it-off than with a common chip model. This evidence contradicts claims that using a common chip model or encouraging students to think about negative numbers as objects makes initial learning of integer arithmetic easier (Küchemann, 1981; Liebeck, 1990). Those who advocate for the benefits of a chip model in order to replicate students' development of whole number ideas, in effect, ask students to hurdle the very conceptual obstacle found in the historical development of negative number arithmetic. This study offers evidence to suggest that such sequencing should avoid replicating the Western historical development of negative number arithmetic by introducing integer arithmetic to students with a collecting objects meaning only after students develop integer arithmetic understanding with a moving-on-a-path metaphor. Research is needed to investigate sequencing and connecting models that draw on different conceptual metaphors or language-based instructional practices that evoke conceptual metaphors (e.g., stories).

More broadly, this study offered insights toward resolving questions of how different integer models could help or hinder learning by comparing individual models and investigating more aspects of integer learning than prior studies, including all four primary operations. These findings might also better translate to predicting research results at scale, because it involved the full range of student achievement found in classrooms and used experimental methods to investigate individual student's thinking and learning after instruction.

Limitations and Future Research

Some of the study design decisions created limitations that leave open questions for future research. These include the length of the instructional unit, researcher-teacher instruction, disentangling impacts of metaphor from model-movements, exclusive focus on two particular models with acontextual instruction, and the particular population studied.

Due to site-based constraints, the unit was limited to eight lessons, about which a student who learned with chips said, "That wasn't enough [time]" when she turned in her tests. When I asked her during an interview how much more time she would recommend, she said one more week, because their units usually took at least three weeks. Even this recommendation, however, is much shorter than the five-week long unit Stephan and Akyuz (2012) reported that they instructed students on only addition and subtraction operations. As did Stephan and Akyuz (2012) and Liebeck (1990), I served dual roles as both researcher and a teacher implementing study instruction. Like these prior studies, the researcher-teachers had extensive teaching experience. Unlike these prior studies, I randomly assigned classes to instruction and used a sample size eight times that of these prior studies. Nevertheless, all such studies that employ a researcher-teacher design could have unintentional bias for which, due to logistical constraints described in methods, the study design did not account. Thus, future replications at scale to understand what students learn in a typical length unit with teachers teaching their own students are necessary and will require a much larger sample size and hierarchical linear modeling to account for different teachers and districts.

The study reported here used a number line model designed to encourage students to move in ways that represent opposite operators. These model-movements differ from other approaches to using a number line, so the findings of this study should not be generalized to other number line models or chip models that use different model-movements. In order to compare models that do evoke different conceptual metaphors, this study did not attempt to separate learning effects due to students' model-movements from effects due to evoking a conceptual metaphor. Future research could address this limitation by assessing impact on learning due to consistency of model-movements for models based on the same metaphor, such as comparing a typical number line model to the walk-it-off model.

The decision to avoid real-life contexts limits the scope of the study implications for integer knowledge development. Future directions for this research that address this limitation include investigating how context interacts with these types of models. Future studies may find that a chip model is beneficial when applied to contexts that draw on a collecting objects metaphor, because these real-life contexts actually promote model-movements consistent with the world (e.g., the ways electrons and protons interact). Beyond performance with context problems, with regard to the process of learning with number lines, a potential limitation of the walk-it-off model should be investigated. As previously noted, some argue that movements in number line models, for example, are arbitrary conventions (Heefer, 2011; Stephan & Akyuz, 2012). To empirically consider such potential limitations qualitative research that investigates whether students themselves view the symbol meanings promoted in the walk-it-off model as arbitrary or in what ways they make sense would be needed. In addition, studies that compare affordances and constraints of a chip and/or walk-it-off model with algebraic approaches are needed to empirically resolve competing claims that models or algebraic approaches are most useful (Heefer, 2011; Vig, Murray, & Star, 2014; Vlassis, 2008). Such research may be particularly important since both models may interfere with disciplinary expectations if students pursue advanced mathematics where people conceive of mathematical objects differently than in school mathematics, such as the plane moving instead of students or chips. Since this aspect of advanced mathematics contradicts human experience in which humans and objects appear to move in the real world, investigations of models that fit such a mathematical conception would require virtual tools.

Practical Implications and Limitations

As discussed in the review of literature, little is known about individual integer models. Thus, this studied focused on comparing student learning with a single integer model in order to better isolate potential model differences as a research design decision. I do not claim that a single model will support students in developing the rich conceptions we desire. This was an additional reason that I conducted this study the year prior to each district’s scheduled integer arithmetic instruction, so that this study would avoid limiting participants' experience. There are several collecting objects, measuring, and moving-along-a-path metaphor-based models that could be experimentally compared to further consider how learning with each model affords and constrains integer learning. Research is also needed to investigate effective sequencing and connecting multiple instructional models. Yet, teachers require the best recommendations to inform how they teach their current students. By the time additional data about integer models is amassed, current middle grade students could be graduating from high school or even college. As a first instructional model, the walk-it-off model was more effective overall for an economically diverse sample (45% free and reduced lunch), improved learning more for those who have less integer knowledge prior to instruction, and draws on the metaphor students most likely already use for negative numbers. While additional research emerges that advances or contradicts the findings reported here, when teachers first introduce integer arithmetic, the motion-aligned walk-it-off number line model should be the most productive model to meet the diverse range of learning needs in real classrooms. Yet, these students were primarily European American fifth- and sixth-grade students in rural districts, so the study should be replicated with other populations to ensure that these results appropriately inform instruction for all students. Anecdotal evidence suggests the walk-it-off model, which a teacher developed and has shared with hundreds of other teachers, is a feasible and low-cost model for teachers to implement (Nurnberger-Haag, 2007). Nevertheless, future investigations should confirm that students in a variety of contexts using these models with typical classroom teachers experience similar results.

Methodological Implications

The delayed posttest results reflect longer-term learning, which although rare in educational experiments, is crucial to make claims about educational impact that matters in students' lives. Relying on immediate testing as most educational studies do could lead to inaccurate conclusions and may be one of the reasons scaling up research has been difficult. Teachers and researchers who study student thinking during instruction or immediately after instruction might for example, have concluded that a chip model was more beneficial for some aspect of integer knowledge, yet the longer-term analysis did not support this. Moreover, consistent with Yelon, Ford, and Golden's (2013) assertions, this study demonstrated that in the short term what seemed to be similarities did actually have different longer-term impacts.

Theoretical Implications

Humans are always moving. Research in cognitive science shows that these movements influence what and how we think (Antle, 2013; Glenberg & Kaschak, 2002). Important work about moving to learn mathematics has begun (e.g., Abrahamson & Trninic, 2015; Gerofsky, 2012; Roth & Thom, 2009). Yet, more is needed, and it is crucial that mathematics education research attend to the *ways* that students move due to instructional models, instead of *whether* they move during instruction. This study contributed to this theoretical goal specific to integer arithmetic and demonstrated that it would be important to investigate if motion-aligned models are more productive instructional models for other mathematics topics.

ARTICLE TWO: HOW STUDENTS' INTEGER MODEL-MOVEMENTS GROUND THE MULTIPLE, DIFFICULT MEANINGS OF "-"

Abstract

Students find negative number arithmetic difficult, but the notation poses additional difficulties. Not only does the same visual symbol "-" have different meanings when positioned differently in relation to other symbols (subtraction operation, negative sign, or opposite operation), but students need to understand multiple meanings of "-" in the same position. This study extends prior research about this notation in terms of population studied, method (experimental), and use of conceptual metaphor theory to interpret the notational meanings that students expressed after instruction with one of two integer models —a chip model or a number line model. Regardless of model, students appropriately distinguished instances where "-" indicated a negative sign versus a subtraction operation. This finding suggests that if students have access to these integer models to ground these basic meanings of "-," they may find it easier to learn this notation earlier than previously reported with more abstract methods. Moreover, with regard to the algebraic meaning of "-" as an opposite operation, interviewed students who used a conceptual metaphor to successfully consider this unfamiliar notation used a motion-on-a-path metaphor by turning the opposite direction on a number line for each "-" symbol (i.e., used the walk-it-off model).

How Students' Integer Model-Movements Ground The Multiple, Difficult Meanings Of "-"

Most mathematics beyond middle school requires proficiency with and understanding about all real numbers, including signed numbers. In addition to proficiently calculating all four operations with signed numbers, mathematically competent people need to be able to order these numbers, use number lines, and use the "-" symbol differently based on the context in which it is positioned (Bofferding, 2014; Chiu, 2001; Lakoff & Nunez, 2000; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010; Sfard, 2000; Thompson & Dreyfus, 1988). Students around the world, however, find this knowledge difficult to develop (Altiparmak & Ozdogan, 2010; Gallardo, 2002; Pierson Bishop et al., 2014; Ryan & Williams, 2007; Vlassis, 2008; Warfield & Meier, 2007). Some documented difficulties include accepting that numbers below zero are valid, placing negative numbers on a thermometer or number line, ordering signed numbers, as well as proficiency with subtraction, multiplication and division (Pierson Bishop et al., 2014; Ryan & Williams, 2007; Warfield & Meier, 2007).

Although the ideas of negative numbers and their operations are difficult, the notation creates additional difficulties for students (Vlassis, 2004). These difficulties with negative number notation impact students' algebraic work beyond the grade levels in which students learned integer arithmetic (Vlassis, 2004, 2008). Most negative arithmetic instruction research and practice has only implicitly addressed these negative and positive sign meanings, but students may need explicit attention to these meanings (Bofferding, 2014 & Stephan & Aykuz, 2012). As Sfard (2007) noted with her aptly named title "when the rules of discourse change, but nobody tells you." Students should have the opportunity to explore and make sense of the underlying mathematical ideas these inscriptions record but notational meanings should be communicated. The written notation or significations of negative number meanings are cultural inscriptions, so I argue, like Bofferding (2014) that it is only fair to offer students the shared meanings of these inscriptions. This transparent approach about negative number symbolism would be analogous to teaching whole number symbolism for quantities. For example, children explore how whole number quantities interact, but are explicitly taught to name the oral and written terms for those quantities in the particular culture in which they live (e.g., “five” using English, but “pyat” in Russian). The study instruction I report here explicitly addressed the meanings of the signs relevant within each model enactment as I compared student learning with a chip model to a number line model. The theoretical framework and methods section regarding lessons explains this focus on grounding the symbol meanings with integer models in greater detail.

Difficulties with Negative Notation

Some difficulties related to what Vlassis (2008) called “negativity,” especially when first working with negative numbers are that students may not accept them as valid numbers with which to operate, but only as results to a subtraction problem (Gallardo, 2002). Students have mistaken negative signs for subtraction operations, particularly in multiplication problems (Ryan & Williams, 2007; Vlassis, 2008). In addition, the unary operational meaning “opposite of” is crucial to student understanding and use in algebra (Nurnberger-Haag, 2007, Vlassis, 2008). For example, in situations such as –(-4), –X, or (-a) + a = 0 neither the subtraction meaning nor negative as a number makes disciplinary sense. Yet, students believe any number with a “-“ sign denotes a negative number, so they often incorrectly conclude that –(-4) can only be negative four (Vlassis, 2008). This operational conception to *take the opposite* is necessary to understand not only when a<0 that –a is actually positive number but also the opposite of algebraic expressions, and recognize when the “-“ symbol could have more than one meaning (Vlassis, 2008). Believing that the "-" sign can have just one single meaning is an obstacle noted with eight grade algebra students (Vlassis, 2008).

Numeral notation and other mathematical symbols such as ÷ are also culturally accepted conventions, but many of these symbols students encounter early in school mathematics denote a single mathematical meaning. The discipline of mathematics, however, has many such symbols that require multiple meanings (Sfard, 2000). The symbols used to indicate signed numbers pose challenges for at least two reasons. First, the same symbol "-" or "+" means different things when positioned differently in expressions (e.g., 4 - 5 means subtract, but "-" in -4 + 5 means "negative") (Sfard, 2000; Vlassis, 2004). Second, even the same "-" sign in an expression can require multiple meanings if the context in which it is situated changes as it does when manipulating algebraic equations (e.g., the "-" in 4 – X =8 means subtract, but after subtracting 4 from both sides of the equation, the "-" needs to mean the opposite of X; Vlassis, 2004). In other words, a subtraction operation becomes an opposite operation in these contexts and requires flexibility to think of the signs in such changeable ways (Vlassis, 2004).

Theoretical Framework

Prior Analyses of "-" Notation

A common framework with which to view these symbols categorizes meanings by how many numbers interact with the "-" symbol. The term unary refers to a negative number itself (e.g., -5), binary to refers to subtraction that needs two numbers (e.g., -5 - 6) (Gallardo & Rojano, 1994). Gallardo and Rojano (1994) identified the symmetric meaning separately without reference to the number of numbers, but which I see as another unary meaning, in which a negative symbol acts on a negative symbol or negative number (e.g., -(-5)). Sfard (2000) used semiotics to classify particular uses of these symbol meanings as structural or operational signifiers. Negative signs are structural signifiers, because the “-” is part of the structure or way to write the numeral of a negative number, which is a mathematical object. The operational signifiers indicate signals to operate or do something such as subtract two numbers or take the opposite of a number or expression.

The first use of the “-” symbol students encounter is the binary operational signifier when young children learn to read or write subtraction equations. Students then at various ages depending on out of school or in school experiences, encounter negative numbers or the unary structural signifier. The discipline of mathematics often uses parentheses as grouping symbols to separate an operational “-” from a unary “-” that indicates a negative number. For example –(-4) or X – (-4). Instead of parentheses, particularly in curricular resources, another way of distinguishing the subtraction operations from negative signs is to superscript the unary “-” signs (e.g., $X - {}^{-}4$; Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006; van De Walle, Karp, & Bay-Williams, 2010). The final meaning of “-,” which Vlassis (2008) called the algebraic meaning, a unary operational signifier that means “take the opposite” is often included in middle grades instruction of integer arithmetic.

In the rest of this article I will refer to these three meanings of the “-” sign using terms that as a prior teacher and current teacher educator I believe could communicate with practicing educators with less translation from research than the terms used in prior frameworks, but with the intent that the ideas of these frameworks described above carry on with these terms. I refer to the negative signs and numbers of which negative signs are a part as a Number Structure meaning. The two operational signifiers I will refer to as Subtraction Operation and Opposite Operation meanings.

Embodied Cognition

Many have noted issues of symbolism in mathematics, because mathematical ideas do not have a specific referent in the world, the way arbitrary symbols like the word "apple" can refer to a physical object (Sfard, 2000). In psychological terms the circularity of symbols referring to other symbols without a referent, such as trying to speak in an unknown language by reading definitions for unknown words in the same unknown language, has been called the symbol grounding problem (Harnad, 1990). To investigate if enacting integer models supports students to ground the multiple and difficult meanings of "-" notation requires using these prior frameworks to notice if students express these meanings in formal mathematical language. Other ways of knowing, however, might be missed if relying on these frameworks alone. Integer instructional models encourage students to physically move in ways that treat arithmetic of numbers as though they exist in the world or are grounded. Thus, I draw on two categories of work in embodied cognition, physical movements and conceptual metaphor theory (CMT) (Glenberg, 2010). Research in cognitive science has demonstrated the impact humans' physical body movements have on cognition (Antle, 2011; Barsalou, 2008; Glenberg & Kaschak, 2002; Day & Goldstone, 2011) and more recently for student learning in science or mathematics (Kontra, et al., 2015; Abrahamson & Trninic, 2015). Thus, how students physically move when enacting integer models warrants attention as a factor of learning. By the term model, I mean not only the tools, but also the physical and verbal processes used (due to affordances, constraints, and instructional choices) with those tools (Nurnberger-Haag, 2014).

Integer model arithmetic in terms of embodied cognition. Two common instructional tools for integers use either chips (in which one color represents the quality of a positive quantity and a different color represents the quality of a negative quantity) or a number line (a line on which tick marks indicate number locations) (van De Walle, Karp, & Bay-Williams, 2010). These tools are not actually numbers, so using such tools treats numbers in metaphorical ways that go beyond language (Kilhamn, 2011). Indeed, Lakoff and Nunez (2000) argued that the discipline of mathematics developed and treats arithmetic in metaphorical ways. In terms of CMT, arithmetic can be thought of using grounding metaphors: *object collection metaphor* (i.e., thinking of numbers as quantities of things that can be manipulated), *measuring stick metaphor* (i.e., lengths measured so numbers are though of as ends of lengths, not points), an *object construction metaphor* (in which numbers are thought of "as wholes made of up parts"), or a *motion-along-a-path metaphor* (in which numbers are points along a path or distances one moves along a path; Lakoff & Nunez, 2000, p. 60-65). I along with others (Kilhamn, 2011) use CMT to categorize the ways integer models encourage students to think about integer arithmetic. I use these ideas, however, as a macro-level way to categorize the patterns of potential ways students can move to enact these models and then specify differences between certain models that draw on the same conceptual metaphor. For example, a typical number line model encourages students to use a motion-along-a-path metaphor to ground meanings of negative or positive numbers or subtraction signs by facing a particular direction (backward or forward) and move in particular directions (right/left, up/down, or positive/negative direction) relative to a number line. In contrast, the walk-it-off number line model (Nurnberger-Haag, 2007), although it also uses a motion-along-a-path metaphor and treats numbers as locations the same as a typical number line model, it specifies students move differently. Students enact a walk-it-off model by determining whether to maintain direction (for positive numbers or addition signs) or move the *opposite* direction (for negative signs or subtraction signs). A chip model grounds meanings of negative or positive numbers in terms of the quality or color of the chips as described earlier and grounds mathematical operations as physical movements of putting in or taking out these chips.

Prior work in CMT has referred to grounding metaphors using nouns as though they come into existence (Chiu, 2001; Kilhamn, 2011; Lakoff & Nunez, 2000). Due to combining theoretical perspectives of conceptual metaphors with the idea that students are physically enacting these metaphors, I use the verb forms of the conceptual metaphors (Nurnberger-Haag, 2014): collecting objects, moving on a path, measuring, and constructing objects. With a focus on either chip or number line models I will continue to focus on the two relevant metaphors of collecting objects and moving on a path.

Model meanings of opposite in terms of embodied cognition. Number line models and chip models may ground a Number Structure meaning for opposite numbers. With a chip model, students represent opposite numbers with the same quantity of opposite color chips. With the walk-it-off model, opposite numbers are points that are the same distance from zero on a number line, but on opposite sides of zero. In this sense, both models treat opposite numbers as things or structures, although they use different metaphors to describe and use those things.

The chip model, however, does not ground the sign "-" as an opposite operation. Neither does a traditional number line model in which students and in Thompson & Dreyfus (1988) walk backwards for negative numbers and left for subtraction. In contrast, the walk-it-off model (Nurnberger-Haag, 2007) was specifically designed to foster arithmetic meanings of "-" as well as foster algebraic meanings that would be consistent with -X. A teacher designed this model with the intent to help students ground the meaning of "-" as an opposite operation by turning the opposite direction.

Focus of Study

Prior research on "-" symbol meanings has offered insights about meanings students can develop much younger than typical instruction (e.g., Bofferding, 2014; Gallardo & Romero, 1999; Goldin & Shteingold, 2001) or long after integer arithmetic instruction in algebra courses (e.g., Vlassis, 2004, 2008). This study offers insights about students' notational thinking immediately after integer arithmetic instruction at a slightly younger age (fifth and sixth grade) than typical instruction but before instructing students about the most difficult, algebraic opposite operation meaning of "-." Moreover, the study experimentally investigated whether working with such notation in terms of the metaphors and motions related to a particular chip or number line model helped students ground the three meanings of "-." The larger study is described in more detail in methods. The analysis reported here focused on what meanings of "-" signs students expressed and if these meanings differed by integer model used during instruction. With a goal to uncover how to help all students better learn this difficult notation, the operationalized research questions follow:

RQ1 What meanings of "-" signs did low, medium, or high performing students express? Which of these meanings were productive in the situation? In what, if any ways, did these expressions differ by integer model experienced during instruction?

RQ2 What solutions and reasoning did students convey when they were asked about the meaning of –(–numeral), notation they had not experienced during instruction?

RQ3 Of interviewed students who offered a positive number as a possible answer to the task in RQ2, what reasoning did they express in making sense of this expression?

Method

Overview of Larger Data Set, Lessons, and Measures

The study was a pre-post-delayed post experimental design that compared student learning with either a particular chip model or a particular number line model (walk-it-off). I randomly assigned the eight classes of students in two rural school districts for whom the administrator and current teacher indicated they had not yet experienced integer arithmetic instruction (initial learners) were randomly assigned to learn with one model. Students at Site A were first semester sixth grade students and students at Site B were second semester fifth grade students. This article reports analysis of the post Symbol Interviews implemented to assess the meanings students expressed for the symbol "-." In the larger study from which this data was analyzed, participating students also took two written tests before instruction (pretest), the day after instruction (posttest), and five weeks after instruction (delayed posttest). These tests consisted of a 46-item skill-based Integer Arithmetic Test (IAT) and a seven-item test that asked students to explain and draw the meaning of problems Explain & Draw Test (EDT). Students also took a one-minute timed fact test to serve as a proxy for prior operational achievement with whole numbers. Twelve of these students also participated in extended task-based interviews about integer arithmetic, to be reported elsewhere.

Specifics About Lessons Related to the Symbol Meanings

As the researcher-teacher I instructed all the lessons for each model to ensure parallel instruction and the teacher remained in the classroom for student safety. During each lesson students worked on tasks and played games in assigned groups and had the opportunity to participate in whole-class discussions. Each class experienced parallel instruction with the same tasks and activities, differing only by the integer model used (the physical representations, the language about how to use those representations and corresponding model-movements). The study instruction consisted of eight lessons of about fifty-minutes focused on ordering numbers, opposite numbers, and integer arithmetic with all four primary operations (addition, subtraction, multiplication, and division). Table 15 provides examples of the instructed operations of subtraction and multiplication with each model. Although students did not explore tasks of opposite operations during the lessons, processes students could use with each integer model to use "-" as an opposite operation are also described in Table 15 since such tasks were posed to students in the interviews. During each lesson I helped students explicitly attend to the meanings of "-" and "+" symbols based in the conceptual metaphor of each model, which are described in Table 16 and discussed in more detail in the subsequent sections.

"-" symbol meanings. Instructional approaches focused on grounding the meaning of the "-" symbols in ways that fit the conceptual metaphors enacted with the models. Students explored ideas of opposite numbers during the first and second lessons using chips or number lines and words (i.e. "-4 is the opposite of 4" and "8 is the opposite of -8"), but opposite operator notation was not taught. I chose not to teach this notation in order to avoid what "a template-driven use" in which students could memorize a limited application of the notation without having a reason for the meanings (Sfard, 2000; Vlassis, 2004, p.482).

Chip model meanings. With the chip model, the "-" when positioned as a negative sign signaled what kind of number positive or negative, which in terms of the model means what kind of chip to use for addition and subtraction. The "-" as a subtraction operation signified to take away or take out and "+" to put in. With multiplication and division operations, the students began with a value of zero, which is also the mathematical starting point for whole number products and quotients with objects although students might not have been aware. One of the "-"

symbols signified the same meaning as for addition and subtraction, that is- what kind of and how many chip and the other "-" symbol signified to take out "-" or put in "+."

Walk-it-off model meanings. The Walk-it-off model for addition and subtraction "-" or "+" signs of the first number meant which number or which point on a number line to begin calculations. The same signs as part of the second numeral signified whether to turn the opposite direction "-" or remain facing the same direction "+" as did the operation of subtraction "-" or addition "+." For multiplication and division, just as the chip model begins with a zero value, so did the walk-it-off model. The walk-it-off model, however, represents zero as a position or point on the number line. Then each "-" signified to turn the opposite direction, whereas a "+" sign signified to stay facing the same direction, which was consistent with the meanings used for addition and subtraction.

Data Sources & Analysis of This Report

Interview purpose, procedures, and tasks. Three to five-minute post Symbol Interviews were conducted one or more days after students took the written assessments. The Symbol Interview was designed to provide opportunities to elicit multiple meanings a student might attribute to the "-" symbols and negative numbers when viewed as part of arithmetic tasks. The purpose was to uncover which of the three meanings of "-" from the literature students used and if they offered different meanings for the same symbol. With the exception of the transfer task -(-56), students were not asked to solve the operation task shown, because written tests and the full interview protocol analyzed elsewhere assessed this operational understanding.

To elicit students' meanings for "-," I adapted a whole number place value task used to elicit students' meanings for written digits in particular positions of a numeral (Kamii & Joseph, 1989). The Symbol Interview consisted of the broad categories of tasks: Orally reading the mathematical expression, explaining the meaning of numerals or "-" signs the interviewer circled in front of the student, and calculating the transfer task of -(-56). The primary questions asked were: What does this that I circled mean? And after the student responded, the interviewer asked "Is there anything else that it could mean?" until the student indicated all his or her ideas had been shared. See Appendix C for the interview protocol. To avoid constraining students' movements, both the interviewer and the student stood during the interview as is typical of gesture-based research (Gerofsky, 2010). With this intention for students to be free to move or gesture, the interviewer held cards that s/he showed students, but students did not hold anything or write.

Interviewers. The author (a female with about twenty years mathematics teaching experience who taught all of the study lessons), a male graduate student, and a female prospective teacher conducted the interviews. All interviewers were the same race as most of the study population.

Interview participants. According to the state website, 45% of the students I instructed had free or reduced lunch and were primarily European American. This analysis answers questions related to student's current thinking post-instruction rather than changes in thinking, so here I report the post interview results. Post symbol interview videos were available to analyze for 91 students. Interviews were analyzed using three approaches for each of the research questions. For the first analysis addressing RQ1, I selected 12 students as matched pairs between conditions. RQ2 analyzed all 91 available post symbol interview videos. RQ3 focused on the subset of post symbol interviews in which 20 students offered positive solutions to the –(-56) calculation task.

RQ1 matched pairs. Twelve post instruction student interviews were selected by posttest IAT low, medium, high with the additional consideration due to the experimental study design that between each integer model condition matched pairs were selected to ensure fair comparisons. Like Vlassis' study (2008) and Chiu (2001), a total of 12 students (6 students from each condition) were selected for symbol meaning analysis using stratified matched pair sampling in order to provide fair comparisons and ensure inclusion of students who began as close to the same beginning characteristics as measured by two pre-IAT and fact test and ended with similar integer arithmetic achievement. I was informed by Vlassis' (2008) stratified selection process of low, medium, and high categories. Based on students' posttest IAT scores, I used criteria of low (below 60%), medium (60% to 80%), and high achieving (80% or above). Two pairs of students from each strata, one who learned with chips and one learned with the walk-it-off method, were selected as matched profiles of students. I looked for students with identical pretest IAT scores one in each condition and then identical (or as close as possible) Timed Fact Test scores. Table 17 reports the scores of the matched pair students. Each matched pair had either identical preIAT scores or differed by 1 point. Note that the four selected pairs in the high and low strata differed on the Timed fact test scores by 0 to 2 points. The moderate strata pairs differed by 4 points and 6 respectively with the student who was assigned to the chip model in both cases having the higher Timed Fact Test score.

Symbol interview videos. Answers offered from the 91 symbol interview videos were analyzed to answer the questions specific to calculating a problem that uses opposite operation notation -(-56). The students who included a positive answer as a possible way to interpret -(-56) were selected for qualitative analysis of their responses including the physical motions and oral language used to express their thoughts to determine what reasoning they used to make sense of this unfamiliar notation.

Analysis

The analysis consisted of identifying students' specific language or physical motion behaviors that indicated symbol meanings. Analysis of students' physical motions can reveal spatial understanding not expressed in words (Church & Goldin-Meadow, 1986; Roth, 2001). These were interpreted in terms of the two frameworks: formal language and conceptual metaphor. The formal language coding focused on ways Vlassis (2008) framework interpreted sign meanings and the conceptual metaphor coding I developed in light of conceptual metaphor theory (Lakoff & Nunez, 2000).

Formal language coding. Two of the interviewers, a rater who coded the opposite signifier EDT item and I, created a list of potential things students say when referring to these types of signs and numbers based on having conducted the interviews. These anticipated codes were used to more efficiently recognize and record typical student responses, but any response students offered were recorded. Some of the typical responses included all combinations of synonyms used for numerals, negative signs, subtraction signs, opposite operation signs (e.g., "negative two take away negative four" "negative two minus negative four" etc). I coded all 12 of the videos with partial blindness to the assigned condition given that the instruction occurred one year prior to coding, but with some possibility of remembering who learned in what way. In order to prevent potential interviewer bias of knowing the strata, scores, or integer model of these matched pair students, videos were sequenced for coding using a random number generator. The other interviewer, who was blind to students' integer model and study hypotheses, coded 15% of the 12 videos for reliability, using a random generation tool to select one student from each condition.

Conceptual metaphor coding. The conceptual metaphor coding scheme was developed first by anticipating potential gestures, drawings, and language that might reveal each metaphor. This coding scheme was then tested and further expanded by interviewing mathematics education graduate students to confirm these anticipated codes and uncover additional behaviors integer arithmetic experts might use. The coding scheme was further revised and clarified on written EDT test items to be reported elsewhere. A prospective teacher and the author developed interrater agreement on pilot data of the written tests before coding full interviews with interrater agreement above 90% (to be reported in future articles). This prospective teacher rater who was blind to student condition as well as the study hypotheses then analyzed the Matched Pair Symbol Interview students reported here and the author coded for interrater agreement, also above 90%.

From the documented behaviors relevant to this analysis I interpreted whether students expressed a particular conceptual metaphor: *Collecting Objects* things (if referred to things, objects, or groups), *Moving on a Path* Points (if referred to numbers as points or locations on a path) or Path (moving on a path including distances moved), *measuring* if students referred to static distances as intervals (which none of the matched pair students did, so this will not be discussed further here). The rater noted presences (1) or absence (0) of each behavior for each interview task. In order to capture language and physical movement relationships between students' expressions of metaphors, language behaviors and physical motions/gestures were separately coded (see Appendix B). Language is inherently sequential, so if multiple metaphors were expressed in language, each of these were noted separately. This was also the case if students' different physical motions communicated different conceptual metaphors in sequence (e.g., first a holding gesture immediately followed by pointing to show movement on a path). If, however, students moved in ways that both held objects and moved those objects along a path, for example, this was deemed an Integrated Gesture in which language and physical motions of any metaphor noticed were coded and described as integrated.

Results

I first discuss the meanings of "-" the 12 matched pair students expressed. The chip model and the walk-it-off model both supported students of all achievement strata to develop similar formal mathematical language meanings of "-." The only differences found between integer models or achievement strata related to whether and how students expressed conceptual metaphor-based meanings of "-." I then report if and how 91 students treated –(N) as an operable expression, where N was a negative number, even though this was new notation that students had not experienced during the unit. Again, the primary differences noticed between students who calculated a positive solution to this expression were due to conceptual metaphors expressed. Specifically, students used the operational aspect of the moving-on-a-path metaphor consistent with the walk-it-off model to reason that the task had a positive solution.

RQ1 Meanings Expressed for "-"

This section describes the 12 matched pair students' symbol interview task responses about negative numerals and "-" symbols in subtraction, multiplication, and in unary opposite operation expressions [i.e., -(-56)]. Student responses to "What does this that I circled mean?" when interviewers circled numerals, negative signs, or subtraction signs are discussed together, because these tasks were designed to elicit meanings for the two basic symbol meanings students experienced during instruction: Number Structure and Subtraction Operation. Then student responses about Opposite Operation meanings are discussed. These meaning categories are displayed in the order in which students from elementary school to algebra develop them in the first column of Table 18. Going down the first column, the first meaning students encounter in school for "-" is a subtraction operation with whole numbers, then negative signs indicating negative numbers, and then opposite meanings. The table displays the number of students who offered each of these meaning for circled symbols organized by their IAT Posttest level (high, moderate, or low) and integer model used during instruction (chip or walk-it-off). For example, when the interviewer circled the subtraction sign in the task $^{-}2$ - $^{-}4$, the first data row shows that two students with high posttest performance who learned with chips and one who learned with the walk-it-off model stated the meaning of "-" was "subtract" or "take away." Shaded cells indicate a meaning is invalid in a particular context (i.e., not a desirable response). Zero to two students could populate each cell, because as described in methods, I selected two students from each condition at each achievement strata for this analysis. Zero students in an unshaded cell indicates no students offered that valid potential meaning. If no students provided an invalid meaning for the context, the cell is shaded, but zeros were omitted for easier reading of the table. For example, the subtraction meaning cells are grayed in the columns for which negative signs were circled because we do not want students to confuse a negative sign in this context with a subtraction sign and no students did, so these are blank. In contrast, it would be valid when a negative sign is circled for a student to indicate the number was the opposite of a positive number, so these cells are not shaded and populated with 0, except the 1 because only one high-level performance student who learned with chips referred to each negative number as the opposite of its positive (e.g., -4 is the opposite of 4).

Within these targeted meanings, students' formal language responses attended to in prior research are discussed first and then I describe student expressions of conceptual metaphors (through physical movements and language). Even if students learned with chips during study instruction, most of the matched pair students who expressed a conceptual metaphor expressed a moving-on-a-path metaphor.

Basic "-" meanings: Number structure and subtraction operation. Each of the 12 matched pair students successfully distinguished "-" signs as part of the Number Structure from a Subtraction Operation where appropriate based on the position within subtraction or multiplication expressions and the inner "-" sign of the –(-56) task. By comparing the rows of student strata within each task column in Table 18, note the similarity of students' formal language responses regardless of whether their total IAT achievement status was low, moderate, or high. Looking down the rows of the table to compare Chip and walk columns of each task type reveal little student difference related to integer model experienced during instruction. Some differences for conceptual metaphor expression are noticeable in the table, but the following qualitative explanations offer greater insights.

Formal language. Not only did all the students distinguish negative numerals from subtraction signs referring to them as "negative numbers" or by name "negative two", but they also expressed understanding that the "-" sign was part of negative numerals and not a numeral itself. This was the case when asked about negative numeral notation or negative signs on all the subtraction and also multiplication operation tasks (see Numeral and Negative Sign columns in Table 18). Note in the table only one person referred to a negative number as the opposite of a particular number. She referred to these as numbers, however, not in terms of an operation with numbers. For example, although –4 can be described with the word "opposite" as the "opposite of four," because the student referred to this as a formal number or a location on a number line, this revealed a Number Structure meaning rather than something to do, or an operational meaning.

Conceptual metaphor based communication. As discussed in the methods, all students experienced integer models that promoted ways of thinking about negative number arithmetic using a conceptual metaphor. During the symbol interviews when asked to read mathematical expressions out loud and explain the meaning of circled numerals and "-" signs, none of the low performing students expressed a conceptual metaphor. Half of the 12 matched pair students expressed at least one of the coded conceptual metaphors (collecting objects, moving on a path, and measuring) and half of the students expressed none of these.

Some students who learned with chips did express a moving-on-a-path metaphor, but none of the students who learned with the walk-it-off model expressed a collecting objects metaphor. One chip-assigned and three walk-it-off-assigned students exclusively expressed a moving-on-a-path metaphor and only two chip-assigned students exclusively expressed collecting objects. One student who learned with chips expressed both a collecting objects and moving-on-a-path metaphor separately depending on the contexts and also as an integrated conception (moving objects down a path).

Collecting objects. Students' verbal explanations and physical motions on these tasks rarely communicated meanings of numerals that indicated conceptions of negative numbers as objects or things. One student who learned with chips, Christie was the only student to explain the subtraction sign as taking away things, which she did both in words and with her gestures. "[Right hand holding gesture] To subtract, which means you would take away [while speaking right held gesture moves away from her body] from it." When asked for another meaning she gestured more strongly and explicitly with both hands to show the boundaries of a quantity as she said "take like what you have from that number and then like [left hand drops and right hand pushes the objects forward away from her body as she says] put aside sometimes"

Moving on a path. Students expressed moving-on-a-path metaphor based meanings to distinguish between a "-" symbol that is just a part of a negative numeral and a negative number. For example, two students who learned with chips and one student who learned with the walk-it-off model referred to the circled negative numerals as points a particular distance below zero (i.e., when -2 was circled, they said "two below zero"). These three students understood that the sign indicated a general location of some point or number on a number line, but in order to know the exact location, one needed the other part of the number as well. When the interviewer circled only the "-" symbol of -5, for example, the students only said that the symbol meant "below zero." They consistently expressed this "number as location" aspect of a moving-on-a-path metaphor for all "-" signs that were part of the Number Structure including -56 in the expression –(-56).

In contrast, Wallace understood that the first number in a subtraction expression was a point on a number line, and accurately referred to the second number after the subtraction sign as an operation or something to do: "that's how far you walk…you turn back around because of the negative sign and then you walk four." He did not, however, offer a moving-on-a-path meaning for the numbers in multiplication expressions. For the subtraction sign he did express a moving-on-a-path operational meaning that "you turn 'cause of the subtraction sign." Another student, Will consistently expressed an operational meaning of "-" both when this symbol was part of a numeral and when positioned as a subtraction sign: "turn the opposite direction."

Integrated conceptions. None of the students who learned with the walk-it-off model integrated a collecting objects metaphor with the moving-on-a-path metaphor meanings promoted during instruction. Just one of the students who learned with chips integrated a collecting objects metaphor into her conceptions of negative numbers due to instruction with chips. Christie used gestures but not language that indicated a conception that the “-“ sign of negative numbers indicated things as she cupped her hand in open ways to hold these things. Each time she did this, however, her gestures integrated holding with moving along a path or turning this held quantity upside down to go “below zero” as she described in language.

Algebraic “-” meaning: Opposite operation. As described in the methods, the –(-56) prompts were used to probe students’ ability to interpret meaning of “-” as an operation other than subtraction when it was positioned next to a negative number. Regardless of integer model used during instruction, students used formal language in similar ways to read the expression out loud and describe the outer “-” sign as negative or part of the Number Structure. Using a lens of conceptual metaphor theory revealed some differences due to integer model when students communicated the meaning of the outer “-”, but not when students simply read the expressions out loud.

Formal language. When asked to read aloud the expression –(-56), most students said: “negative, negative fifty-six.” None of the students on this transfer task used the phrase “opposite of.” This reading of the expression was consistent with the meaning they offered when the interviewer circled the outer “-” and asked “What does this that I circled mean?” Although all students hesitated, sensing that the outer “-“ should be different from the inner “-“ when encouraged to offer some meaning, the only meaning they could attribute was the same meaning as the inner “-.” As shown in Table 18, 11 of the 12 students said the outer “-“ was a “negative” implicitly meaning a negative sign or explicitly stating “negative sign.” Only one student talked about “-” as a negative number not a sign. Thus, 11 of 12 students distinguished the “-” sign from a negative numeral even on this transfer task. Although one student Caleb did refer to the outer “-“ as subtraction, the first meaning he expressed was a negative sign. This student, however, was the only student to provide inaccurate additional ways to read mathematical expressions (e.g., claiming that $^{-}2$ - $^{-}4$ could also be read as “two minus four”).

Conceptual metaphor based communication.

Collecting objects. As previously described, when asked about the meaning of the inner “-” sign student Christie expressed an integrated gesture; however, for the outer sign she still gestured as if holding the quantity -56, but did not move it along a number line. Her gesture and her language indicated she did not know what to *do* with the quantity, because as she said, [that outer sign] “it’s the symbol for a negative number and it’s by itself right now.”

Moving on a path. When asked about the outer sign “-” students Chloe and Warren said it meant “below zero.” Notice below zero represents a position, which is a Number Structure meaning, not an operation needed for this outer “-“ symbol. Will, however, offered the same additional operational meaning he did for all “-“ symbols in subtraction, multiplication, and –(-56) tasks: “It means turn the opposite direction.”

RQ2: Do Students Treat –(N) as An Operable Expression, Where N Is a Negative Number?

To address the second research question about how students interpret a notational situation of –(N), where N is a negative number, first I report all student solutions for the –(-56) calculation interview task, “If you had to calculate an answer to this problem, what do you think the answer might be?” (n=91). I then address the third research question by analyzing in more detail all the students who offered reasoning to support a positive solution on this task (n=20). This unary calculation task was unfamiliar to students, because except the case when students did subtraction problems (i.e., binary operations), the unit did not provide experience with the notation of two “-“ signs next to each other. Thus, most students initially responded to the interview prompt with uncertainty or “I don’t know.” When the interviewer encouraged students, however, with a prompt such as “It’s okay if you don’t know, but what do you think it might be?” 82 students provided at least one potential answer or reasoning (even if they did not explicitly state a numerical solution). The solutions these students offered included negative solutions (-157, -112, -59, -56, -54), zero, and positive solutions (1, 23, 56, 58, 112, 250, and whatever 56 × 56 or -56 × -56 would be) as well as statements such as “greater than -56,” it would be “negative” or “positive.” In order of frequency, the four most common solutions offered were “negative fifty-six” (n=36), “fifty-six” (n=20), “zero” (n= 13-14) and “negative one hundred twelve” (n= 6).

Students answered -56 most often, because the students thought any “-” that was not an operation of subtraction indicated a negative number. Students explained that there was no operation to perform, so what was written there was negative fifty-six, so in their minds it had to remain negative fifty-six. Although they said words like “*stay* negative fifty-six [emphasis added],” as though the numeral would not be changed, when asked “Would it be written the same as it is now?” students said they would rewrite the solution without the additional negative sign. Students determined zero to be a solution in various ways that usually involved treating the outer “-“ sign as subtraction by imagining a number that was not written on the task card, such as 56 or 0 (i.e., 56 – (-56). Students commonly provided four of these answers, because they interpreted the two negative signs to signify to do something two times. The most common of these were to multiply by the number 2 or to multiply either 56 or -56 twice (i.e., 56 × 56, -56 × -56, -112 or 112). The solutions such as negative fifty-nine or negative fifty-four seemed to have simply been misstatements when orally reading negative fifty-six. Some students provided solutions such as “fifty-six” but the reasoning was not explicit whether they knew the solution was positive. If students used multiplicative reasoning to explain such ideas, I used their written test solutions described in detail in other manuscripts (see Article One; specifically, their written skill-based negative number multiplication and division scores on the IAT and an EDT item that asked them to explain the solution to the product of two negative numbers) to interpret if students likely intended their solution to be positive. These students were included in the next analysis of positive solutions. The students for whom I could not confirm that they intended a positive solution were excluded (4 students who learned with chips and 4 who learned with the walk-it-off model).

RQ3 Symbol Interview Students Who Posited Positive Solutions

I share the explanations that supported the accurate solution of 56 or other positive answers, because even if inaccurate, this indicates a crucial conceptual shift. That is, students were able to overcome the idea that when no binary operations are present, a "-" can only indicate a negative number (Number Structure). Table 19 displays the number of students who learned with each integer model who used each type of reasoning. As the totals show, twice as many students who learned with the walk-it-off than chip model determined a positive solution. Students from both conditions used generalized rules or ideas of turning the opposite direction on a number line to make sense of the -(-56) as equal to 56 or some other positive number. Regardless of whether students grounded the meaning in a moving on a path metaphor or used a rule, most of these students used multiplication to make sense of –(-56) as an operable expression with a positive solution. It seemed to be students' experiences with parenthetical notation as one way to write multiplication in prior mathematical experiences and during the study instruction that inspired them to read this notation using multiplication. Notice in the table that both chips and walk students used generalized rules and none of the students from either integer model used a collecting objects metaphor. Moreover, only students who learned with the walk-it-off model used the operational aspect of the moving-on-a-path metaphor in ways that led to a positive solution. More students who answered 56 used the meaning of "-" as something to do, specifically to turn the opposite direction on a number line, than a generalized rule (see Table 19).

"-" Can only have meaning in relation to another. Ten students offered explanations to justify that -(-56) would be positive by using knowledge about the product of two negative numbers or stated rules about notational symbols interacting (i.e., relationship of two negative signs situated next to each other). Table 19 provides examples of these reasoning categories of Generalized Solutions, Generalized Notation, and Notational Mnemonic. Note that students who learned with either model expressed Generalized Solutions, which students extrapolated from having calculated and participated in generalizing activities about multiplication problems during instruction. Evidence suggested that the one student who used a trick for remembering that two sequential negative signs should become positive (notational mnemonic), had been taught this prior to the study instruction.

Only one of these ten students who used generalized language to reason about positive solutions communicated conceptual metaphors about these ideas in words or physical movements. This student had learned with chips, but her gestures conveyed a meaning of the opposites as locations on a number line (Number Structure). Most students kept their hands in a resting position during the entire task and a few students pointed to signs on the task card, but did not communicate spatial meanings with their gestures. In spite of having learned to multiply integers with one of two conceptual metaphor-based models, even those students who calculated the answer of –(-56) as a multiplication problem of two negative numbers (Generalized Solutions reasoning) did not express conceptual metaphors in words or through movement.

"-" Has its own meaning moving on a path: Turning the opposite direction. As Table 19 indicated, 10 students used a meaning of the "-" sign grounded on a number line as "turning" or "opposite" to determine a positive solution, seven of whom determined the correct answer 56. Eight of these students who learned with the walk-it-off model used turning the opposite direction to multiply numbers in three different ways: turning to multiply, turning to multiply -1 × -56, and turning to multiply by two or twice. Examples of each reasoning type are shared next.

Turning to multiply to get to 56. Table 20 and 21 display two students' responses who treated –(-56) as the multiplication of two numbers by turning the opposite direction for each "-." Table 20 shows Frank who while determining the solution did not express a conceptual metaphor, but then while explaining how he reasoned, gestured in ways consistent with a moving-on-a-path metaphor. Although all of the students' language-based responses in Table 20 and 21 used the term "opposite direction" emphasized during instruction in order to label the movements with the canonical terms to support development of taking an opposite, two other students simply said "turn around." Will first figured out the solution by physically turning his whole body and talking under his breath, before offering the formal response in Table 21 in which he again moved his entire body. Note that all of the students who used these reasoning approaches were able to refer to some numbers as points on a number line (Number Structure understanding), yet also treat "-" notation as an operation (Opposite Operation).

Turning to multiply -1 × -56. In several of the cases just described, students viewed the parentheses as a multiplication operation to which they could then apply their knowledge of "-" meanings for multiplication operations. Winnie extended this idea to the lack of a numeral before the parentheses but a "-" symbol as potentially being an omission of the number -1. Table 22 displays an example of this student who determined *-1* × *-56* would be positive (affect and posture were not noteworthy to include in the table). Winnie thought of this approach to imagine a negative one in front of a mathematical expression on her own. This approach is a common instructional strategy algebra teachers use to help students and a valuable mathematical idea in order for -1 to be a number that distributes across an expression.

Turning to operate twice. Three students made sense of the two negative signs as indicating both to turn the opposite direction and to perform some other operation twice, either multiplying twice or adding twice (multiply by 2). The conversation between the interviewer and Wayne illustrates why students might consider the notation to indicate negative fifty-six times negative fifty-six. This student only moved his hands from a position resting on his hip or collar at one point and maintained similar affect throughout the task.

> Wayne: *Uhmm, probably the answer for negative fifty-six times negative fifty-six.*
>
> I/A: "Ok and why did you think that it would possibly be negative fifty-six times negat (student starts to respond)"
>
> *Wayne: - cause in times parentheses* [Index finger moves in a way that seems to write parentheses in the air] *might mean times [pause] in math*
>
> I/A: "But why, I only see one negative fifty-six there so why did you think that it might be negative fifty-six times negative fifty-six?"
>
> Wayne: *might be a shorter way to just do that if you write just one*
>
> I/A: "Ok, so you're thinking that this extra symbol here might just be a short-cut for saying do it times itself but don't write it all out?"
>
> Wayne: *mmhhmm* [meaning yes]

Although Wayne did not explicitly state the solution to -56 × -56 would be positive, I interpreted it as a positive solution, based on evidence from his written IAT and EDT tests, which also reveals that his conception of integer arithmetic used an operational meaning for "-": (a) he accurately answered all multiplication and division skill problems (b) he explained why the product of two negatives is positive because each “-“ means to turn the opposite direction (c) and he accurately explained that –(negative number) would be a positive number because you would turn the opposite direction twice.

Although in the oral interview Wayne only implied that a “-“ means turn the opposite direction, another student, Rick explicitly stated that the negative signs have an operational meaning on their own: “the *negatives mean* they tell you to turn around…*the other way* [emphasis added].*”* Table 23 shows why Rick thought the solution could be 112 as an example of multiplying by two as well as turning the opposite direction twice to obtain a positive solution. The student’s choice of the phrase “either way” to describe how one should move on a number line to represent “-” illustrates a crucial difference in the way the walk-it-off model promotes moving on a number line compared to other number line models that could be used. Rick was aware that “-” does not tell you whether to move in a particular direction on the number line, but whether to move in the opposite direction from which he is facing, so that same sign could tell him to go in “either” direction depending on the context. Thus, this provides evidence that after using the walk-it-off model to develop a meaning for “-“ his conception of operations with integers was not restricted to thinking of positive numbers moving in a positive direction or negative numbers moving in a negative direction.

***Collecting objects*.** Just as a student who used the walk-it-off model conceived of –(-56) as -1 × -56, a student using chips or any other cancellation model could have thought of this expression this way. To model this process a student would begin with zero and then take out one group of 56 negatives. This process is the same as conceiving of “-“ as 0-(-56), which could also yield an accurate solution. Although –(-56) could be conceived of in these ways, none of the interviewed students used these collecting objects based ideas to determine a positive solution.

Discussion: Physical Movements Ground Notational Meanings

For students with low, moderate, and high overall integer arithmetic achievement, this study provided evidence that using physical experiences to ground meanings of “-” while learning integer arithmetic with either instructional model supported students’ symbol sense with regard to the basic meanings of “-.” This lends support to Vlassis’ (2008) claim that students need a “concrete meaning” for this notation and potential solutions with regard to how ground the symbols (Harnad, 1990) with respect to integer notation. She argued that algebra students had difficulty with the meaning of “-,” because they learned about negative numbers on a number line, but learned integer arithmetic with an abstract rule-based approach. With respect to the most difficult meaning of “-,” the algebraic opposite operation, however, one model did provide better support. In unfamiliar notational contexts, more students used the conceptual metaphor and physical motions consistent with the walk-it-off model to treat “-“ as an operation other than subtraction.

Number Structure: Negative Numbers are Valid

It was noteworthy that each of the conceptual metaphor based models used in this study supported students’ acceptance of negative numbers as valid numbers. Moreover, this was the case for students of all achievement strata. This contrasts with reports of students who often did not accept negative numbers as valid numbers without being the result of a subtraction operation, even though these prior reports were the same age and slightly older than the students in the study reported here (Gallardo, 2002).

Number Structure and Subtraction Operation Meanings of "-"

Compared to prior research with even older students (Gallardo, 2002, Vlassis, 2008), students in the current study expressed better understanding for the basic meanings of the sign "-" (Subtraction Operation and Number Structure). Whereas prior research found students confused negative signs for subtraction signs, particularly within multiplication contexts (Ryan & Williams, 2007; Vlassis, 2008), this study provided evidence that grounding "-" for the primary operations with an appropriate model supported students in avoiding these errors. Both the chip model used in this study and the walk-it-off model helped students infer a subtraction operation meaning for "-" only in its proper location in an expression. This was the case even though the study unit duration was short and regardless of whether students used a model that drew on a collecting objects or moving on a path metaphor. Thus, it seems grounding the meanings of the notation in some appropriate conceptual metaphor supported student development of Number Structure and Subtraction Operation meanings for "-".

Opposite Operation Meanings

Although the different integer models students used in this study did not result in differential student understanding for the basic meanings of "-," the walk-it-off model did better support more students to conceive of "-" as an operation other than subtraction. Algebra students have been critiqued for not conceiving of the "-" as an opposite operation when that meaning is needed (Vlassis, 2008). Yet, the results of this study demonstrated that two aspects of knowledge could support initial learners to successfully expand their understanding of "-" from Number Structure meaning toward an opposite operation meaning: thinking about integer multiplication and a physical grounding of the meaning of "-" as turning the opposite direction on a number line. The students in the current study, however, were younger and had not been taught this notational meaning. In spite of these demographic differences that might predict students in this study would have even more difficulty, 20 students (about one-fifth of the interviewed students) made sense of this unfamiliar notation. Almost three-quarters of these students learned with the walk-it-off model. Moreover, it was the walk-it-off model that offered productive ways for ten students to ground "-" as meaning to do something, or operate, specifically to turn the opposite direction on a number line to represent an Opposite Operation, rather than a Number Structure. It was this operational aspect of the moving-on-a-path metaphor unique to the walk-it-off model that supported student reasoning. Again, this approach to moving on a number line differs from the typical number line models used, but I leave additional discussion for future research that compares typical number line models with a walk-it-off model.

When students from either the chip or walk-it-off model thought of "-" as a point or location on a number line, this structural meaning of number seemed to inhibit their ability to conceive of the outer "-" as an operation to calculate –(-56). This informs us of two implications. First students in upper elementary or middle school may already bring to their explorations of negative numbers a grounded structural conception of numbers as locations on a number line (moving on a path metaphor) that they will likely continue to use even if they learn with chips. Second, whether this meaning is one of the meanings ignored with a chip model or reinforced with a number line model (and thus a potential limitation of a number line model), students may use this negative numbers as points below zero meaning of "-" to their tasks in ways that reinforce a Structure of a Number meaning in an inhibitive way.

"Opposite of" notation was intentionally left as transfer assessment tasks to investigate if one model would better help students infer an operational meaning for the notation without explicit instruction. Through the data I came to recognize that the instructional focus on opposites as numbers in both models (as points in the walk-it-off model and as opposite color chips in the chip model) may have reinforced some students' rigid adherence to a Number Structure meaning that may have inhibited thinking about "-" as an operation. The sequence and standard as written in the Common Core Standards for Mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) might serve this same inhibitory conception, because this document recommends instruction of –(-x) notation as part of a focus on the Number Structures in sixth grade, a year before students operate on negative numbers. In spite of the potential issue I identified with the study instruction, the experiences students had with the walk-it-off model supported more students in conceiving of the "-" as something to do, which is the basis of an operation. Future research could investigate the effect of using these models during lesson activities in which students compare both meanings of the signs—opposite number (Number Structure meaning) and also "-" as something to do (Opposite Operation).

Symbol Sense

The findings that students could offer multiple meanings for a "-" sign and the interpretations of possible meanings for –(-56) notation also demonstrated positive findings about some students' broader understanding of mathematical symbolization. As Vlassis (2008) found, some students believed a "-" sign in a particular context could only have one meaning. Yet, contrary to Vlassis (2008), many matched pair students offered multiple meanings for a "-" sign. This included multiple meanings for the exact same symbol in an expression (e.g., "-" meaning a negative sign indicating a negative number and to turn the opposite direction).

The students who interpreted the expression –(-56) as -56 × 2, -56 × -56, or -1 × -56 also revealed their broader sense of how the discipline uses notation to efficiently communicate mathematical ideas. Although the meanings of -56 × 2, -56 × -56 students offered were not the canonical ways people invented to notate multiplying by two or exponentiation, these students' ideas that a second "-" meant to multiply -56 by itself or add it a second time are reasonable conventions, which the people who developed this notation could have chosen. The historical mathematicians who developed the arbitrary notation that current people accept, could have chosen the same notation these fifth and sixth grade students reasonably interpreted. For example, although the canonical disciplinary notation for -56 × -56 is $(-56)^2$, the fifth grade student 3000 had not yet learned about exponential notation, but like mathematicians in history offered that "it might just be faster to write it [-56] once." Some students' ideas that two signs means operate two times are actually insightful. The canonical mathematical interpretation could mean do the same operation twice. Although the expression could be interpreted in the most common way as a single operation to take the opposite of the structural signifier -56, it can also be thought of as two opposite operations (take the opposite of 56 and then take the opposite again).

Limitations and Future Research

In Article One that included delayed post written tests, I argued for the need for longer-term studies. Thus, I must acknowledge that, in spite of logistical constraints, a lack of delayed interviews is a limitation of the study reported here. Future research would benefit from incorporating follow-up interviews. If interview results follow similar patterns as the written test results in Article One, then even the lack of differences at post-interview such as on formal language meanings expressed for "-" could actually become different weeks later.

Withholding instructional experience with opposite operation notation in order to treat these as transfer tasks created a limitation that did not offer low performing students sufficient opportunity to make sense of this notation. Given the findings here and that the walk-it-off model was originally designed to assist algebra students who struggled with mathematics or language in general (Nurnberger-Haag, 2007), research that includes this notation in instructional activities may find it better supports students who might typically struggle more with mathematics to develop the algebraically productive conceptions of "-" as an Opposite Operation.

Influencing Curricular Placement and Implications for Algebra

In order for students to successfully conceive of "-" as an operation, we should ensure that instruction supports both Number Structure and Opposite Operation meanings. Two suggestions for curricular sequencing of negative number arithmetic arise from this study. First, teaching negative number arithmetic with related notation by fifth or sixth grade is feasible. Second, opposite operation notation should be taught after students learn integer operations. The fact that at least some moderate and higher achievement students in fifth and sixth grade were individually able to overcome thinking of "-" as strictly part of a number to successfully interpret the –(-56) transfer offers additional evidence that cooperative group instruction with negative number arithmetic need not wait until later middle school. Such evidence contradicts curricular placement as in Belgium where teachers are not allowed to instruct students on negative number arithmetic before age 12 (Heefer, 2011). In terms of curricular documents in the United States, these pieces of evidence indicate that future research should investigate moving integer arithmetic instruction from seventh grade to at least as early as sixth grade, if not earlier.

Second, -(n) where n is a negative number, notation should be taught after students learn all four primary operations. Several pieces of evidence and theory support this assertion. Given the field advocates building on student thinking and the results suggested that multiplication was a resource for students' understanding "-" as an opposite operation, multiplication should be taught before opposite operation notation. Moreover, this curricular recommendation is consistent with theoretical recommendations that understanding numbers can only be understood through operating on them, so instructing students on the number structure meaning of opposites separate from the operational meaning by a year or more could cause unneccesary obstacles for students by separating the ability for students to work with mathematical processes and objects together (Sfard, 1991). Students likely need opportunities to compare Number Structure and Opposite Operation meanings of an opposite value at the time they are introduced.

Vlassis noted that in addition to notational difficulties, the difficulty two eighth grade students had solving an equation involving a product was partly due to “the impossibility of giving a concrete meaning to the product of” negative numbers (2008, p. 566). The findings of this study and (Nurnberger-Haag, Article One) that grounding meanings for integer multiplication and division by turning the opposite direction on a number line not only supported understanding integer products but could lead to understanding “-” as an opposite operation. Future research is needed to assess the algebraic implications after students’ use the walk-it-off model for initial integer learning.

Theoretical and Methodological Implications

Prior research regarding negative number arithmetic including that which attended to conceptual metaphors privileged language based communication of “-” notation and negative number arithmetic (e.g., Chiu, 2001; Vlassis, 2004, 2008). Yet, this study demonstrated, as others have that spatial understanding is a valid way of understanding and is not always translated into language (Alibali, Evans, Hostetter, Ryan, & Mainela-Arnold, 2009). In fact, those who express spatial understanding are often the higher performing (Gerofsky, 2010; Sassenberg & Van der Meer, 2010). Moreover, this study like (Church & Goldin-Meadow, 1986) suggested that knowledge as it is developing is even more likely to be expressed spatially than verbally. Thus, if we truly want to assess student conceptions and not just what students can express in words, our research methodology must include ways of accessing spatial understanding as this study contributed.

APPENDICES

APPENDIX A: Tables and Figures

Table 1

Mapping integer models in terms of related theoretical perspectives of conceptual metaphors

	Cancellation Chip Model[1]	Number Line: Typical Number Line	Number Line: Walk-it-off model
Mathematical Object or Process Meanings	Collecting Objects	Moving on a Path	Moving on a Path
Numbers	Quantities of objects	• Points or positions on a number line • Distance to move & which direction to face	• Points or positions on a number line • Distance to move
Negative or positive signs	Kind or quality of object (negatives or reds; positives or blacks)	• Part of a number that signals the point is below or left of zero • Face left or negative (face right or positive)	• Part of a number that signals the point or position is below or left of zero • Move or turn the opposite direction on a number line
Operations			
Addition	Put in objects	Move right or up (positive direction)	Move in same direction
Subtraction	Take out objects	Move left or down (negative direction)	Move in opposite direction
Multiplication	Group objects	Repeated distances facing forward or backward	Turn/maintain direction then move product of absolute values
Division	Group objects	Repeated distances facing forward or backward	Turn/maintain direction then move quotient of absolute values
Opposite Of	NA	NA	Turn the opposite direction

Table 2

Examples of Integer Model-Movement for three types of operations

Operation	Cancellation Chip Model[1]	Number Line: Typical Number Line	Number Line: Walk-it-off model
Subtraction -3 - -2=-1	• Put in 3 negatives (reds) • Take out 2 negatives (reds) • 1 negative remains	• (-3) Stand on the point • (-2) Face negative • (Subtract) Move backwards	• (-3) Stand on the point facing positive direction • (Subtract) Turn the opposite direction • (- of -2) Turn the opposite direction again • Move 2 in the direction facing
-3 – 2=-5	• Put in 3 negatives (red) • Put in at least 2 negatives (red) and 2 positives (black) • Remove 2 positives (black)	• (-3) Stand on the point • (2) Face positive • (Subtract) Move backwards	• (-3) Stand on the point facing positive direction • (Subtract) Turn the opposite direction • (+ of 2) Maintain direction • Move 2 in the direction facing
Multiplication -2 * 3=-6	• Start with 0 as nothing • Put in 3 groups of 2 negatives (reds) • 6 negatives (red)	• Start at the point 0 • (3) Face positives • (-2) Move backward two sets of three spaces	• Start at the point 0 • (- of -2) Turn the opposite direction • (+ of 3) Maintain direction • Move 6 (product of 2 * 3)
-2 * -3=6	• Start with 0 as nothing • Use multiple chips to represent 0: Put in at least 3 groups of 2 negatives (reds); Put in the same number of positives (blacks) • (-3) Take out 3 groups of 2 negatives (reds) • 6 positives (blacks) remain	• Start at the point 0 • (-3) Face negatives • (-2) Move backward two sets of three spaces	• Start at the point 0 • (- of -2) Turn the opposite direction • (- of 3) Turn the opposite direction • Move 6 (product of 2 * 3)

Table 2 (Cont'd)

Opposite Of	NA	NA	-(-7) • Start at the point 0 • Outer "-" Turn the opposite direction • Inner "-" of -7 Turn the opposite direction • Move 7

[1] This chip model is the common chip model that uses separate chips of two different colors (e.g., black and white chips) and encourages students to put in extra chips with a value of zero only when needed.
[2] Another typical number line model has students start at 0 on the number line then move to the first number of the problem, just as some chip models have students start every problem by representing zero with chips.

Table 3

Lesson sequence with variations due to model of task wording or activities

Lesson	Topic	Purpose	Main Lesson Activities and Tasks	Chip Variations	Walk-it-off Variations
1	Introduction Opposite Numbers Extending Numbers to negative numbers by subtracting and adding	• Introduction, establishing norms, assigned trios/pairs	• Exploring opposites • Use addition and subtraction problems to motivate a need for negative numbers.		Walk-it-off model students also learned to construct number lines
2	Addition & Subtraction	• Explore addition and subtracting with negative numbers.	• In trios/pairs, calculate given problems • [1]Represent -4 (In trios/pairs write 3 equations)	Write…for the value -4.	• Constructed number lines • Write… to arrive at the number -4 on a number line.

Table 3 (Cont'd)

3	Addition & Subtraction Ordering Numbers Generalizing about Real Numbers	• Become proficient with adding and subtracting negative numbers • Understand when a negative number is less than or least, greater than or greatest • Realize that subtracting a number does not always make a smaller value	• Represent 0 (In trios/pairs write 3 equations) • Greatest & Least number game (Decide whether to add or subtract rolled numbers.)	• Represent "value of 0"	• Constructed number lines • Represent "how to arrive at the number 0"
4	Multiplication & Division	• Encourage students to attend to "-" signs • Compare processes for addition and subtraction versus multiplication and division	• Use model to multiply and divide whole numbers then extend the model to negative numbers	NAV	NAV
5	Multiplication & Division	Explore multiplication & division with negative numbers	• Multiplication and division problems in trios and pairs	NAV	NAV

Table 3 (Cont'd)

6	Multiplication & Division	Become proficient with multiplication and division	• Multiplication and division problems in trios and pairs • [2]Individual students make conjectures on exit tickets in response to	NAV	NAV
7	All primary operations		• [2]Trios/pairs determine if conjectures always sometimes or never true • Individual mixed review then discuss with trio/pair	NAV	NAV
8	All primary operations	• Review how to calculate with all primary operations • To generalize about operations	• [2]Finish conjecture discussions • Individual mixed review then discuss with trio/pair	NAV	NAV

Note. NAV=No additional variations in planned tasks beyond differences in how students represent ideas with each model as described in Table 2.

[1] Modified from CMP in which students use chips to create the value -2 (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006).

[2] Modified from Heck & DeFord (2012)

Table 4

EDT Generalization Item Purpose and Examples

Item	Purpose: Understanding Assessed	Examples
Opposite Sums	Generalization that sums of opposites equal zero	Trina and Jaleesa are students in your grade at another school. Trina said that $^-8 + (^-7 + 7)$ does <u>not</u> give the same answer as $^-8 + (^-5 + 5)$. Jaleesa said they will. Circle who is right: Trina or Jaleesa. Draw and write an explanation in words to convince a friend that this student is right.
Generalizing Operations	Generalizations of the four primary operations	Look at each of the four choices a, b, c, and d. Each rectangle hides the **same negative integer**. Which answer choice makes the greatest number? (Write the letter here). _____ Use words and drawings to convince the students that you are right. a) 4 + ■ b) 4 × ■ c) 4 − ■ d) 4 ÷ ■
Opposite Operator	Generalization of opposite operator meaning for the symbol "-"	A classmate stood at the board and said, "I am thinking of a negative number. The classmate wrote on the board: My secret number is (negative number) Then the classmate wrote: Now my secret number is $^-$-(negative number) **Even though it is impossible to figure out the exact secret number, what CAN you tell the students at the other school about the green secret number your classmate is thinking of now?** Explain **how** you know.

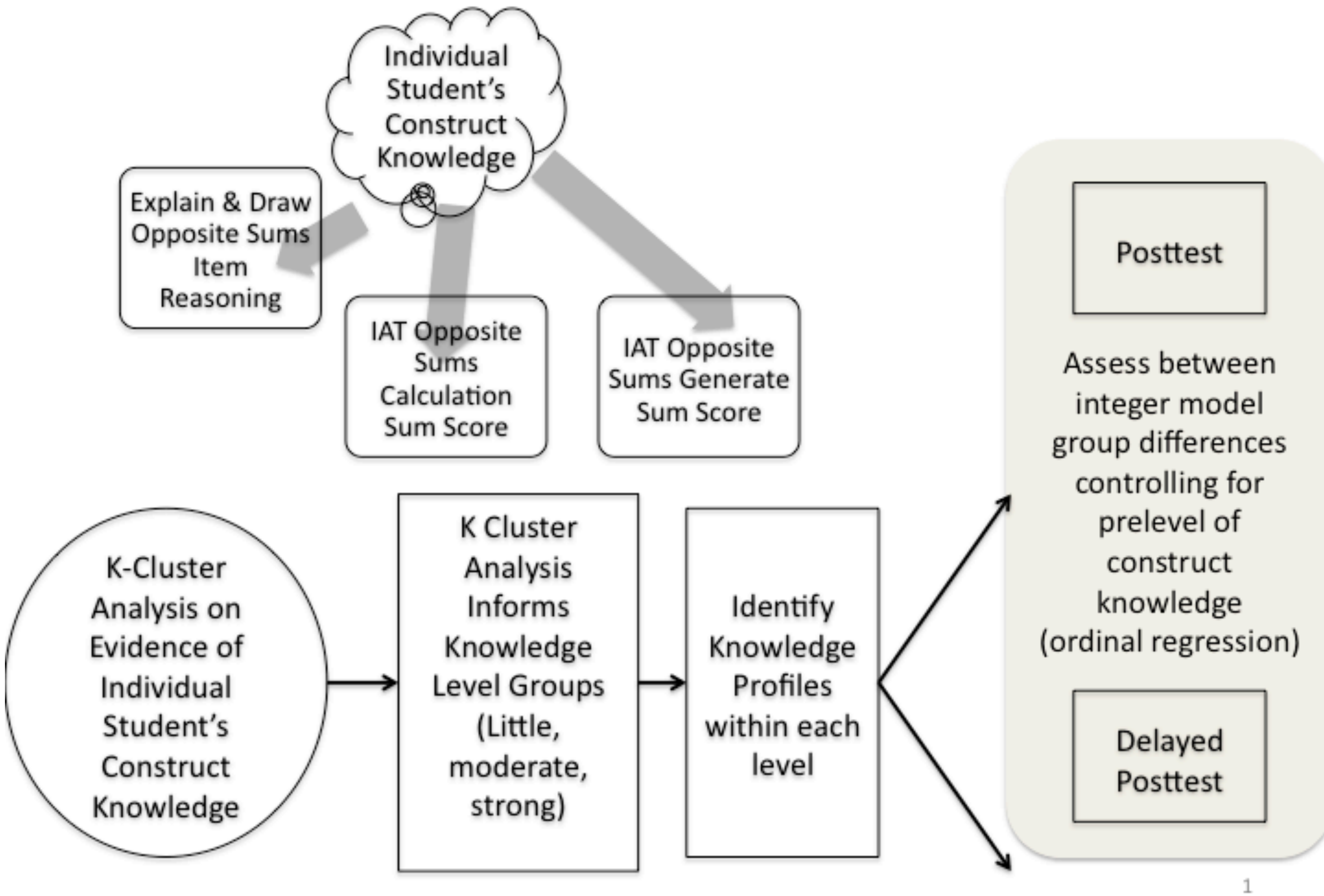

Figure 1. Mixed method analysis process for combining qualitative item reasoning with quantitative scores to determine evidence of an individual's opposite sum knowledge before testing for between integer model differences.

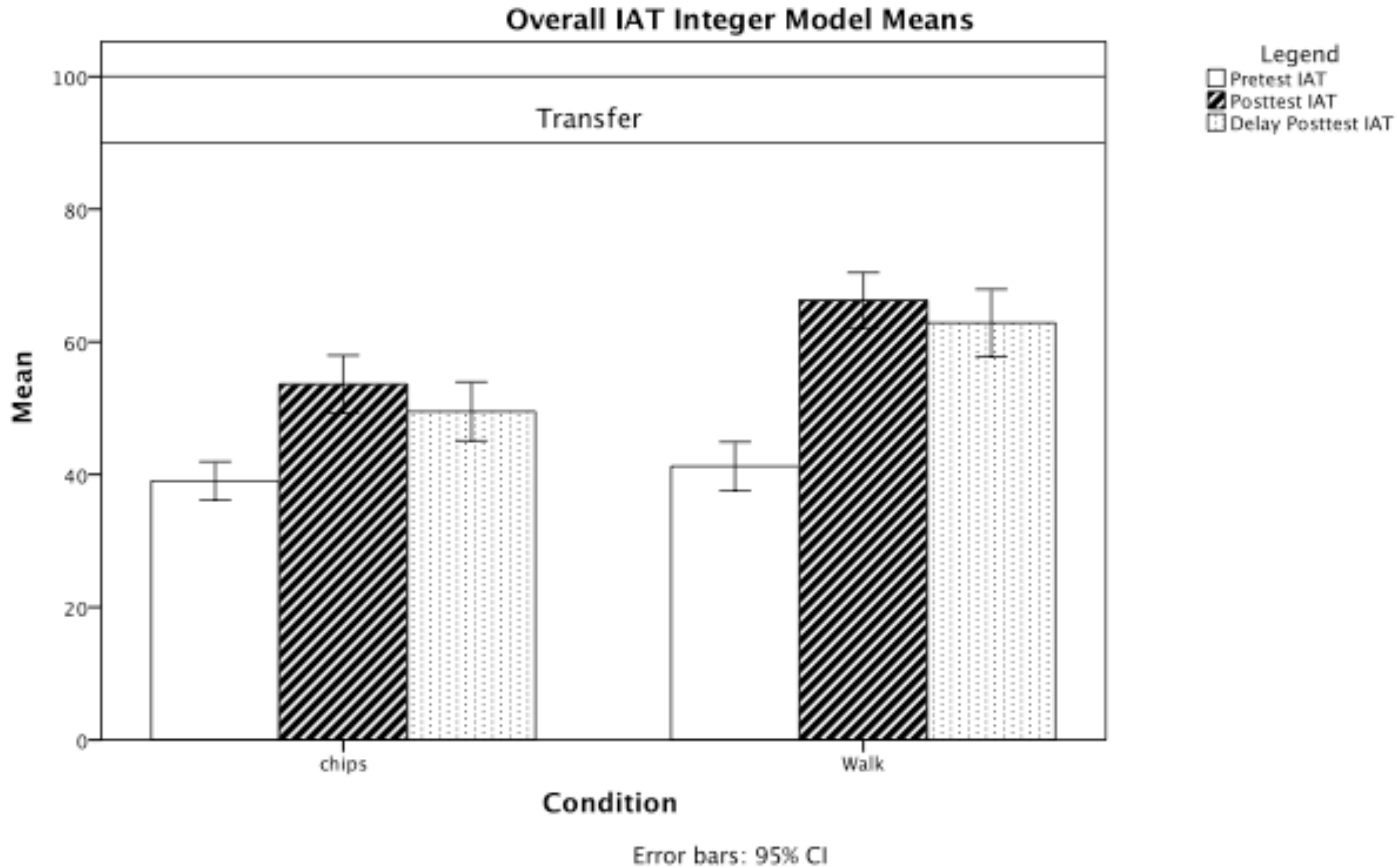

Figure 2. Bar graphs of pretest posttest and delayed posttest total IAT unadjusted means by integer model condition.

Table 5

Overall IAT Test ANCOVA Results

	Posttest[a]			Delayed Posttest[b]		
	F	p	np²	F	P	np²
Integer Model	17.1	< .001***	.101	18.8	< .001***	.112
Pretest IAT	53.3	< .001***	.260	43.1	< .001***	.225
Fact Test	1.65	.201	.011	2.61	.108	.017
Gender	7.61	.007**	.048	5.01	.027*	.033
MD Preconceptions	12.1	.001**	.073	5.09	.026*	.033
District	3.04	.083	.020	.351	.554	.002

Note.

Df for all post predictors (6, 152); delayed post (6, 149)

* p< .05, ** p< .01 ***p <0.001

Table 6

Overall IAT Test ANCOVA Parameter Estimates

	Posttest			Delayed Posttest		
	β	95% CI[,]	SE	β	95% CI	SE
Integer Model	-10.4***	[-15.3, -5.41]	*2.50*	-12.3***	[-17.9, -6.71]	2.84
Pretest IAT	.713***	[.52, .91]	*.098*	.711***	[.49, .93]	.108
Fact Test	.257	[-.14, .65]	*.200*	.369	[-.08, .82]	.228
Gender	6.79**	[1.93, 11.7]	*2.46*	6.23*	[.73, 11.7]	2.78
MD Preconcept	-9.20**	[-14.4, -3.97]	*2.65*	-6.74*	[-12.7, -.84]	2.99
District	4.52	[-.60, 9.64]	*2.59*	1.74	[-4.06, 7.54]	1.74

Note. Adj. R^2 : a=. 458, b=.422

* p< .05, ** p< .01 ***p <0.001

Table 7
MANCOVA Multivariate Tests Inconsistent vs. Consistent Model-Movements

	Posttest[a]			Delayed Posttest[b]		
	F	p	np^2	F	P	np^2
Integer Model	28.62	< .001***	.275	28.61	<.001***	.280
Pre Consistent	9.73	< .001***	.114	13.49	< .001***	.155
Pre Inconsistent	4.91	.009**	.061	4.71	.010**	.060
Fact Test	3.07	.049*	.039	5.02	.008**	.064
Gender	3.17	.045*	.040	2.35	.099**	.031
MD Preconceptions	4.82	.009**	.060	3.15	.046*	.041
District	1.11	.334	.014	.288	.750	.004

Note. Df for all post predictors (2, 151); delayed post (2, 147)
* alpha .05, ** alpha .01 ***alpha <0.001

Table 8

MANCOVA Multivariate Tests Consistent v. Inconsistent Parameter Estimates Posttest and Delayed Posttest Raw Scores

	Posttest						Delayed Posttest					
	Inconsistent[a]			Consistent[b]			Inconsistent[c]			Consistent[d]		
Variable	β	CI[,]	SE	β	CI[,]	SE	β	CI[,]	SE	β	CI[,]	SE
Integer Model	-3.24***	[-4.22, -2.26]	.72	-.843	[-2.26, .57]	-6.6	-3.54***	[-4.56, -2.51]	.52	-.897	[-2.58, .61]	.80
Pre Consistent	.222***	[.11, .33]	.08	.314***	[.16, .47]	4.0	.251***	[.14, .37]	.06	.431***	[.25, .61]	.09
Pre Inconsistent	.389**	[.14, .63]	.18	.394*	[.04, .75]	3.1	.398**	[.14, .65]	.130	.370	[-.03, .77]	.20
Fact Test	.083*	[.01, .16]	.06	.129*	[.02, .24]	2.2	.121	[.04, .20]	.04	.161*	[.04, .28]	.06
Gender	1.06*	[.11, 2.02]	.70	1.63*	[.26, 3.01]	2.2	.578	[-.42, 1.58]	.51	1.71*	[.15, 3.26]	.79
MD Preconcept	1.46**	[.41, 2.52]	.77	2.22**	[.70, 3.74]	2.7	1.26*	[.15, 2.37]	.56	1.96*	[.23, 3.68]	.87
District	.290	[-.71, 1.29]	.73	1.05	[-.40, 2.49]	.57	-.029	[-1.08, 1.02]	.53	.487	[-1.15, 2.12]	.83

Note. Adj. R2 a=.418 b=.265, c=.438, d = .4288

* alpha .05, ** alpha .01 ***alpha <0.001

Table 9

MANCOVAs Multivariate Tests Degree of Consistency Model-Movements Posttests and Delayed Posttests Standardized Scores

Variable	Posttest F	p	np^2	Delayed Posttest F	P	np^2
Integer Model	20.6	< .001***	.295	20.6	< .001***	.301
Pre Consistent	6.31	< .001***	.113	8.65	< .001***	.153
Pre Inconsistent In Process	5.11	.002**	.094	3.67	.014*	.071
Pre Inconsistent Beginning	.315	.814	.006	.384	.765	.008
Fact Test	1.66	.179	.032	4.34	.006**	.083
Gender	1.71	.168	.033	1.78	.153	.036
MD Preconceptions	2.02	.114	.039	1.65	.181	.033
District	1.26	.289	.025	.398	.754	.008

Note. Df for all post predictors (3, 148); delayed post (3, 144)
* alpha <.05, ** alpha <.01 ***alpha <0.001

Table 10

MANCOVA Degree of Model-Movement Consistency Parameter Estimates of Posttest Standardized Scores

	Inconsistent Beginning[a]			Inconsistent In-Process[b]			Consistent[c]		
Variable	β	CI[,]	SE	β	CI[,]	SE	β	CI[,]	SE
Integer Model	-.862***	-1.11, -.61	.127	-.414**	-.71, -.12	.151	-.144	-.42, .14	.141
Pre Consistent	.263***	.13, .40	.067	.142	-.02, .30	.079	.283***	.14, .43	.074
Pre Inconsistent In Process	.203**	.07, .33	.066	.223**	.07, .38	.078	.238*	.09, .38	.073
Pre Inconsistent Beginning	.049	-.09, .19	.069	-.002	-.16, .16	.082	-.011	-.16, .14	.077
Fact Test	.014	-.01, .03	.010	.021	-.002, .05	.012	.021*	-.001, .04	.011
Gender	.249*	.002, .50	.125	.026	-2.7, .32	.148	.262	-.01, .54	.139
MD Preconceptions	.268	-.01, .55	.141	.271	-.06, .60	.167	.346	.04, .66	.156
District	.107	-.15, .37	.132	.171	-.14, .48	.156	.273	-.02, .56	.273

Note. Adj. R2 a=.426 b=.188, c=.290, * p <.05, **p <.01 ***p<0.001

Table 11

MANCOVA Degree of Model-Movement Consistency Parameter Estimates of Delay Posttest Standardized Scores

	Inconsistent Beginning			Inconsistent In Process			Consistent		
Variable	β	CI[,]	SE	β	CI[,]	SE	β	CI[,]	SE
Integer Model	-.926***	-1.18, -.68	.127	-.386**	-.67, -.10	.145	-.163	-.44, .11	.140
Pre Consistent	.257***	.13, .39	.066	.235**	.08, .38	.076	.345	.20, .49	.073
Pre Inconsistent In Process	.202**	.07, .33	.006	.150*	.001, .30	.075	.187*	.04, .33	.073
Pre Inconsistent Beginning	.061	-.07, .20	.069	.029	-.13, .19	.079	-.004	-.15, .15	.076
Fact Test	.017	-.003, .04	.010	.040	.02, .06	.011	.025*	.003, .05	.011
Gender	.148	-.10, .40	.125	-.088	-.37, .20	.143	.238	-.03, .51	.138
MD Preconceptions	.282*	.003, .56	.141	.097	-.22, .41	.161	.283	-.02, .59	.155
District	.062	-.20, .40	.132	.004	-.29, .30	.150	.148	-.14, .44	.145

Note. Adj. R2 a=.443 b=.253, c=.308

*p <.05, ** p<.01 ***p <0.001

Table 12

Performance for Ordering Integers Pre to Post and Pre to Delayed Changes

	Pre-Post		Pre-Delayed	
	Chips	Walk	Chips	Walk
Category	No. %	No. %	No. %	No. %
Improved	30 (36.1)	31 (39.2)	28 (35.0)	28 (35.9)
Maintained	37 (44.6)	40 (50.6)	35 (43.8)	41 (52.6)
Regressed	16 (19.3)	8 (10.1)	17 (21.3)	9 (11.5)
Total	83	79	80	78

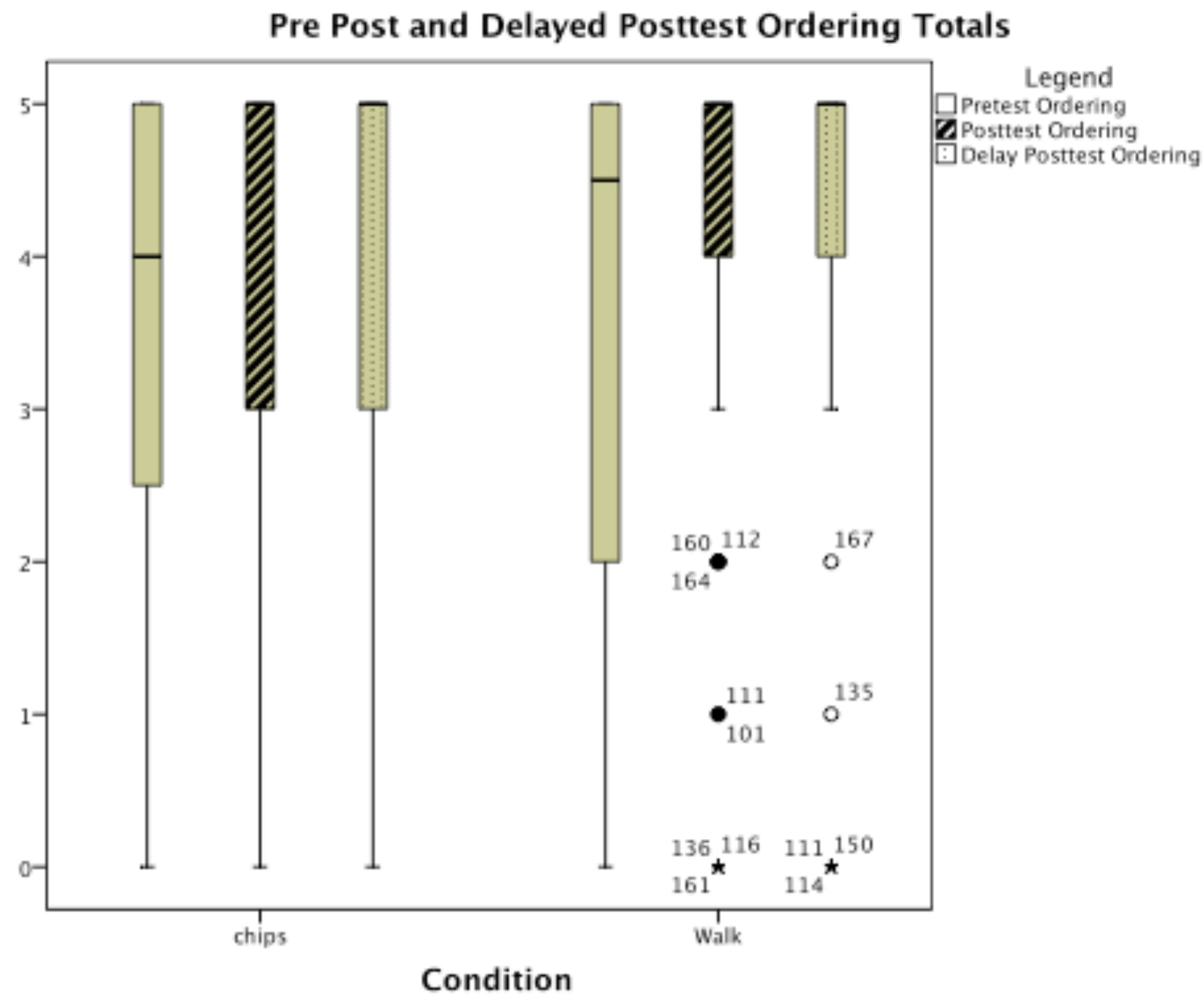

Figure 3. Box plots of pretest posttest and delayed posttest ordering number totals by integer model condition.

Table 13

Conceptual metaphors expressed on EDT pretests for ordering item or single digit operation items

	Collecting Objects		Measuring		Moving on a Path	
	No.	%	No.	%	No.	%
Ordering Numbers	7	11.1	2	3.2	40	63.5
Operations	9	14.3	0	0	13	20.6

Note. Four intervals of randomly assigned student identification numbers selected for analysis (n= 63).

Table 14

Summary of integer model benefits by aspect of integer knowledge

Aspect	Posttest	Delayed Posttest
Ordering Numbers	n.s.	Walk*
Operations		
Sums of Additive Inverses	Chip**	n.s.
Generalizing Operations Accuracy	n.s.	Walk*
Generalizing Operations Reasoning	Walk***	Walk***
Primary Operations by degree of consistency of model-movements	Walk***	Walk***
"-" as meaning "opposite of" [Transfer]	UA	UA

Note: *** p<.001, **p<.01, * p<.05, n.s.=not statistically significant using α<.05; UA= unable to analyze

Table 15

Examples of Integer Model-Movement for three types of operations

Operation	Cancellation Chip Model[1]	Number Line: Typical Number Line	Number Line: Walk-it-off model
Subtraction -3 - -2=-1	• Put in 3 negatives (reds) • Take out 2 negatives (reds) • 1 negative remains	• (-3) Stand on the point • (-2) Face negative • (Subtract) Move backwards	• (-3) Stand on the point facing positive direction • (Subtract) Turn the opposite direction • (- of -2) Turn the opposite direction again • Move 2 in the direction facing
-3 – 2=-5	• Put in 3 negatives (red) • Put in at least 2 negatives (red) and 2 positives (black) • Remove 2 positives (black)	• (-3) Stand on the point • (2) Face positive • (Subtract) Move backwards	• (-3) Stand on the point facing positive direction • (Subtract) Turn the opposite direction • (+ of 2) Maintain direction • Move 2 in the direction facing
Multiplication -2 * 3=-6	• Start with 0 as nothing • Put in 3 groups of 2 negatives (reds) • 6 negatives (red)	• Start at the point 0 • (3) Face positives • (-2) Move backward two sets of three spaces	• Start at the point 0 • (- of -2) Turn the opposite direction • (+ of 3) Maintain direction • Move 6 (product of 2 * 3)

Table 15 (Cont'd)

-2 * -3=6	• Start with 0 as nothing • Use multiple chips to represent 0: Put in at least 3 groups of 2 negatives (reds); Put in the same number of positives (blacks) • (-3) Take out 3 groups of 2 negatives (reds) • 6 positives (blacks) remain	• Start at the point 0 • (-3) Face negatives • (-2) Move backward two sets of three spaces	• Start at the point 0 • (- of -2) Turn the opposite direction • (- of 3) Turn the opposite direction • Move 6 (product of 2 * 3)
Opposite Of	NA	NA	-(-7) • Start at the point 0 • Outer "-" Turn the opposite direction • Inner "-" of -7 Turn the opposite direction • Move 7

[1] This chip model is the common chip model that uses separate chips of two different colors (e.g., black and white chips) and encourages students to put in extra chips with a value of zero only when needed.
[2] Another typical number line model has students start at 0 on the number line then move to the first number of the problem, just as some chip models have students start every problem by representing zero with chips.

Table 16
Mappings of models to "-" meanings

Meanings	**Chip Model**	**Walk-it-off Model**
Number Structure		
Numbers	Quantities of objects	Points or positions on a number line
"-" (Negative Signs)	Kind or quality of object (negatives or color)	• Part of a number that signals the point or position is below zero or • Move or turn the opposite direction on a number line
Subtraction Operation	Take away or remove objects	Move or turn the opposite direction on a number line
Opposite Operation	NA	Move or turn the opposite direction on a number line

Table 17
Symbol Interview Matched Pair Scores

		Chip			Walk		
Post IAT Strata	Pair	Student	Pre IAT	Fact	Student	Pre IAT	Fact
High	1	Chloe	52	25	Will	53	24
	2		38	13	Warren	39	15
Moderate	1	Caleb	38	25	Wallace	39	21
	2	Christie	46	18		45	12
Low	1		32	21		32	16
	2		21	15		21	14

Note. Pseudonyms are given for those students named in text for additional analysis.

Table 18

Matched Pair Symbol Interview "-" Meanings expressed when interviewer circled numerals, negative signs or opposite signs

Meanings			Level	Symbol Circled in Expression									
Category	Specific Response	Examples	IAT Level	Subtract Sign		Numeral		Negative Sign		Inner "-" -(-56)		Outer "-" – (-56)	
				C	W	C	W	C	W	C	W	C	W
Subtraction Operation	Subtract	"Subtract" "Take away"	H	2	1								
			M	2	2	1							
			L	1	2							1	
Number Structure	Negative Number	"It's a negative number" "Negative two"	H			2	1						1
			M			2	1	1					
			L			2	2						
	Negative Sign	"Negative [sign]" "It's [a sign that means] negative" "It means the number is negative"[1]	H					2	2	2	2	2	1
			M					2	2	2	2	2	2
			L					2	2	2	2	2	1
	Opposite Number	"The [number that is] opposite of four"	H			1	0	1	0	1	0	1	0
			M			0	0	0	0	0	0	0	0
			L			0	0	0	0	0	0	0	0
	Number below zero	Two below zero It's below zero	H			1	1	1	1	1	1	1	1
			M			1	0	1	0	1	0		
			L			0	0	0	0	0	0		

Table 18 (Cont'd)

Opposite Operation	Move the opposite direction	Turn the opposite direction	H	0	1	0	0	0	0	0	1	0	1
			M	0	1	0	0	0	1	0	0	0	0
			L	0	0	0	0	0	0	0	0	0	0

Note. IAT Level= IAT posttest level high, moderate, or low performing; C=chip model student, W=Walk-it-off model student

[1] "minus" was coded for as a synonym of "negative" and also as a synonym for "subtract." The student in this case used the synonym for negative.

[2] Student said negative number but other evidence….

Table 19

Symbol Interview Calculated –(-56) Confirmed Positive Solutions

Reasoning	Description and Examples	Chip	Walk
Moving-on-a-path metaphor	"-" means turn the opposite direction on a number line	0	10
Collecting Objects metaphor	"-" means take out 56 negatives or take out one group of 56 negatives	0	0
Generalized Rule			
Generalized Solutions	Products of two negative numbers have a positive solution *Fifty-six, er negative fifty-six.* [Interviewer/author "Which one?"] *well there's two negative signs so I would think a positive fifty-six. Um well with what we've learned we've learned that cause like the parentheses mean multiplication and then whenever you have two negative signs you get a positive answer so it would be positive fifty-six.*	1	2
Generalized Notation	Two "-" signs interact to become positive *Fifty-six. Because I see two negative signs and I'm thinking that they would cancel each other out, leaving the fifty-six.*	5	1
Notational Mnemonic	Memory trick about the form of the signs *Fifty-six. ah because if you might ah, because if you take one of that and you take that and you turn it sideways it equals a positive sign.*	0	1
	Total	6	14

Table 20

Student (Frank) example of gesturing to communicate turning to multiply to get to 56

Noteworthy affect or posture	Physical motions	Speech
		Um
[6 seconds thinking in a relaxed posture leaning against the wall with a furrowed brow]		
	bolts to standing	*Oh!*
Firm voice	Points for emphasis	*Fifty-six*
	Points to card	*Because there's two multiplication*
Giggles		*er not multiplication ah negative signs, and if there's one negative sign*
	Points left index finger down	*that means you start at negative*
	Turns left index finger to point up	*and if there's another one that means you go positive. You turn the opposite direction.*

Table 21

Student (Will) example of moving whole body to reason and also communicate turning to multiply to get to 56

Noteworthy affect or posture	Physical motions	Speech
	Turned entire body with elbows bent 90 degrees at sides hands in blades pointing forward	*Regular fifty-six. Positive fifty-six. Because you face positive and you always start on zero if it's times or division.*
	Keeping arms in the same positions, turned entire body the opposite direction	*Then the negative number means turn the opposite direction and you go up whatever it is"*

Table 22

Student (Winnie) example turning to multiply -1 × -56

Physical motions	Speech
Points to signs on the card	*Negative fifty-six. Because I'm just thinking that there would be a one there or something. and imagine if there were parentheses right there and multiply one- negative one by negative fifty-six.*
Cracks fingers and stretches while I1 talks. Nods	I1: "So negative one times negative fifty-six is fifty-six [pause] or negative fifty-six? I1: I can't remember what you said." *Fifty-six.*

Table 23

Student (Rick) example turning to multiply by 2 to obtain a positive solution

Noteworthy affect or posture	Physical motions	Speech
Both hands rest in sweatshirt pocket for the duration of the interview unless noted.		*I'm not sure, um. About like one hundred twelve. Because fifty-six times two is somewhere around one hundred twelve.* [I1: "Why do you think its negative fifty-six times two?] c*ause it's in parentheses* *I think that the negatives mean —they tell you to turn around on the number line either way.* I1: "you said one hundred twelve did you mean positive one hundred and twelve or did you mean positive or negative or?" *umm looks like positive because if I think you can um*
Student emphasized intonation bolded	points to card then hands go behind his back.	*That* *means to start on zero and* ***face*** *the negatives*
Student emphasized intonation bolded	left index finger points right (the positive direction on a horizontal number line)	*and then you turn around again and you would face the positives.* *So* ***positive*** *one hundred twelve.*

APPENDIX B: Conceptual Metaphor Coding Definitions

Table 24
Abbreviated Coding Definitions for Conceptual Metaphors

Metaphor	Behavior/Coding Definition	
	Words	**Gesture/Drawings**
Moving on a Path	PTS-W: Numbers as points or locations on a path (e.g., "This is my zero" or "Here is 3" "It" refers to positions or points on a path)	PTS-DG: Pointing, flat hand, etc. that are deictic gestures or drawings to show numbers as positions. In a drawing more likely students draw a point, circle positions, draw arrow to it, etc.
	PATH-W: Speech that indicates use of a number line with motion (up, down, left, right, forward, backward, opposite, turn, toward, away from, etc.)	PATH-DG: Gestures or drawings (arrows, loops/jumps, etc.) that indicate motion ideas on a **number line**
	TURN-W: If uses specific language of turn or opposite, then also code this code.	TURN-DG: Turning head side to side, rotating or turning hand, indication of rotation or turning with a single finger, turning body core (engagement of spine), flipping, reflecting, etc. In a drawing representations of flipping, reflecting, or circling around.
	DISTANCE-PATH-W: Language that refers to distance but using motion and underlying conception of numbers as positions not measuring stick	DISTANCE-PATH-DG: Gesture or diagrams that show distance from body to another point, etc. indicating endpoints of a distance achieved by moving to compare based on accompanying speech, directional arrows on a drawing

Table 24 (Cont'd)

Collecting Objects	THINGS-W: Reference to chips, things, or specific objects (e.g., candy, bags or groups implies things) in words.	THINGS-DG: Drawings or gestures (acting out) of chips, things, or specific objects
	MANIP-W: Words that describe the ways object manipulatives were moved for integer operations during instruction (see training on how objects are manipulated in problems).	MANIP-DG: Gestures or drawings that model the ways object manipulatives were moved for integer operations during instruction (see training on how objects are manipulated in problems).
	HOLD-W: Words that indicate holding, grasping, supporting (e.g., cupped hand, pinching fingers as though grasping, flat hand facing up as though supporting or holding something, things in a bowl, etc.)	HOLD-DG: Gestures or drawings that indicate holding, grasping, supporting (e.g., cupped hand, pinching fingers as though grasping, flat hand facing up as though supporting or holding something, things in a bowl, etc.)
	TOUCH-W: Words that indicate touching separate things at once	TOUCH-DG: Gestures that indicate touching separate things at once (three fingers as though the three fingers represent or are touching three different objects, etc.)
	COUNT-W: Words that indicate counting discrete objects (If counting on number line, do not code here.)	COUNT-DG: Gestures that indicate counting discrete objects (e.g., pointing in sequence that may or may not be lined up, may be accompanied by verbal explanations that indicate counting to distinguish from pointing in a way that indicates location along a path. If counting on number line code below not here.)
Measuring Stick	STICK-W: Language that refers to distance as a static measurement	DISTANCE-STICK-DG: Distance between two points indicated by an interval (e.g., interval C-gesture (see Williams, 2009) or two hands, fingers, bars on a line segment
	COMPARE-W: Comparison of line segments or intervals to find the distance or lay on top of the other (so motion is due to comparing lengths, not to find a given length).	COMPARE-DG: Comparison of line segments or intervals to find the distance or lay on top of the other (so motion is due to comparing lengths, not to find a given length).

APPENDIX C: Symbol Interview Protocol

Set Up

If it's okay with you, I'd like us both to stand so that you feel free to move around or do whatever you think would help you communicate how you are thinking. Is that okay?

Is it okay if I set up the video camera now, but I won't start recording yet? [Make sure that the entire person's body is visible in the view even if he/she takes a few steps either direction]

Do you have any questions or comments? [Pause]

Is it okay with you to do a sound check? You can just say testing 1 2 3. [pause for confirmation].

I'll start the video now. [Start the video. Make sure that the entire person's body is visible in the view even if he/she takes a few steps either direction]

[Playback the video, if the student's voice is clear, then say], ok. *We have sound, so we're all set. Is it okay if I start the video for real now?*

Interview Begins

[Be sure to have the back up audio recording]

Every time I ask you a question, I'll ask if there is another way until you tell me there isn't another way or you can't think of one, because I want to make sure you have a chance to say all of your ideas. Ok?

So I'll just keep asking you if there's another way, until you tell me there isn't or you can't think of one.

(Show the student the card)

B) Ex: - 2 - (-4)

1) *Would you please read out loud what is written on this card?*

2) *Are there any other ways that you can say what is written on this card?*

3) *Is there another way?*

(**Circle -2** with a blue pen)

What does this that I circled mean?

(Point to the circle, not the -2)

(Paraphrase what the student said, *So it means….?*

Pause for student confirmation.

Is there anything else it can mean)?

(Show student a new card with the same problem)

This card says the same thing, but doesn't have my pen marks on it.

(**Circle −4**)

What does this that I circled mean?

Paraphrase *So it means….?*

Pause for student confirmation.

Is there anything else it can mean?

(Show student a new card with the same problem)

This card says the same thing.

(Circle **subtraction sign**)

What does this that I circled mean?

Thanks. Let's look at a card with a different problem.

[Show the student the card.]

Ex: - 3 - -5

1) *Would you please read <u>this</u> card out loud?*

2) *Are there any other ways that you can say what is written on this card?*

3) *Is there another way?*

(Circle − sign of −5)
What does this that I circled mean?
Is there any thing else it can mean?

This card says the same thing, but doesn't have my pen marks on it.
(Circle − sign of −3)
What does this that I circled mean?
And is there anything else it can mean?

C) -5 × -7

1) *Now would you please read **this** card out loud?*
2) *Are there any other ways that you can say this?*
3) *Is there another way?*

Thanks. What does this that I circled mean?
(Circle −7)
Is there anything else it means)?

This card says the same thing, but doesn't have my pen marks on it.
(Circle −5)
What does this that I circled mean?
Is there anything else that it means?

D) -4 × -8

1) *Would you please read **this** card out loud?*
2) *Are there any other ways that you can say this?*
3) *Is there another way?*

This card says the same thing, but doesn't have my pen marks on it.
(Circle **− sign** of −4)
What does this that I circled mean?
Is there anything else that it means?

This card says the same thing, but doesn't have my pen marks on it.
(Circle − sign of −8)
What does this that I circled mean?
Is there anything else that it means?

E) -(-56)

1) *Would you please read **this** card out loud?*

2) *Are there any other ways that you can say this?*

3) *Is there another way?*

This card says the same thing, but doesn't have my pen marks on it.
(Circle **− sign** of −56)
What does this that I circled mean?
Is there anything else that it means?

This card says the same thing, but doesn't have my pen marks on it.
(Circle − sign outside of the parentheses)
What does this that I circled mean?
Is there anything else that it means?

If you had to calculate an answer to this problem. What do you think the answer would be?
Paraphrase, *So it means…?*
Pause for confirmation.

REFERENCES

REFERENCES

Alibali, M.W., Evans, J.L., Hostetter, A.B., Ryan, K., & Mainela-Arnold, E. (2009). Gesture-speech integration in narrative: Are children less redundant than adults? *Gesture, 9*, 290-311. doi: 10.1075/gest.9.3.02ali

Abrahamson, D. (2009). Embodied design: constructing means for constructing meaning. *Educational Studies in Mathematics, 70*, 27-47. doi: 10.1007/s10649-008-9137-1

Abrahamson, D, & Trninic, D. (2015). Bringing forth mathematical concepts: signifying sensorimotor enactment in fields of promoted action. *ZDM: Mathematics Education, 47*, 295-306. doi: 10.1007/s11858-014-0620-0

Altiparmak, K., & Özdoğan, E. (2010). A study on the teaching of the concept of negative numbers. *International Journal of Mathematical Education in Science and Technology, 41*, 31-47.

Antle, A. N. (2011). Balancing justice: Comparing whole body and controller-based interaction for an abstract domain. *International Journal of Arts and Technology*.

Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between theoretical and conceptual work. *For the learning of mathematics, 7*, 245-274

Ball, D. L. (1993). With an eye on the mathematical horizon. *The Elementary School Journal, 93*(4).

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Berlinghoff, W. P., & Gouvêa, F. O. (2002). Something less than nothing?: Negative numbers Math through the ages: A gentle history for teachers *and others.* (pp. 81-86). Farmington, MA: Oxton House Publishers

Bofferding, Laura (2014). Negative integer understanding: Characterizing first graders' mental models. *Journal for Research in Mathematics Education, 45*, 194-245.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2*, 155-192.

Bruno, A., & Martinon, A. (1999). The teaching of numerical extensions: The case of negative numbers. . *International Journal of Mathematical Education in Science and Technology, 30*, 789-809.

Chiu, M. M. (2001). Using metaphors to understand and solve arithmetic problems: Novice and experts working with negative numbers. *Mathematical Thinking and Learning, 3*(2), 93-124.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43-71.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*.

Cotter, S. (1969). Charged particles: A model for teaching operations with directed numbers. Arithmetic Teacher, 16(5), 349-353.

Creswell, J.W. (2008). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 3rd Edition. Los Angeles: SAGE.

Day, S. B., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 551-567.

de Freitas, E., & Sinclair, N. (2012). Diagram, gesture, agency: theorizing embodiment in the mathematics classroom. *Educational Studies in Mathematics, 80*, 133-152.

Duke, N.K., & Beck, S.W. (1999). Education should consider alternative formats for the dissertation. *Educational Researcher, 28*(3), 31-36.

Ekstrom, R.B., French, J.W., Harman, H.H., & Dermen, D. (1976). *Manual for Kit of Factor-Referenced Cognitive Tests* Educational Testing Service.

Ernest, P. (1985). The number line as a teaching aid. *Educational Studies in Mathematics, 16*(4), 411-424.

Fennema, E., Franke, M.L., Carpenter, T.P., & Carey, D.A. (1993). Using children's mathematical knowledge in instruction. *American Educational Research Journal, 30*, 555-583.

Fischbein, E. (1987). Intuition in Science and Mathematics: An Educational Approach. Boston: D Reidel Publishing Company.

French, D. (2001). Two minuses make a plus. *Mathematics in School, 30*(4), 32-33.

Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dordrecht, Holland: D, Reidel.

Fuson, K.C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction, 7*(4), 343-403.

Gallardo, A. (2002). The extension of the natural-number domain to the integers in the transition from arithmetic to algebra. *Educational Studies in Mathematics, 49*(2), 171-192.

Gallardo, A., & Rojano, T. (1994). School Algebra. Syntactic difficulties in the operativity. In D.

Kirshner (Ed.), *Proceedings of the Sixteenth International Conference for the Psychology of Mathematics Education: North American Chapter* (pp. 159-165). Baton Rouge, LA.

Gallardo, A., & Romero, M. (1999). Identification of difficulties in addition and subtraction of integers in the number line. *Proceedings of the twenty-first International Conference for the Psychology of Mathematics Education: North American Chapter* (1, 275-282). Cuernavaca, Morelos, México.

Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 109-128). New York, NY: Cambridge University Press.

Gerofsky, Susan. (2010). Mathematical learning and gesture: Character viewpoint and observer viewpoint in students' gestured graphs of functions. *Gesture, 10*, 321-343. doi: 10.1075/gest.10.2-3.10ger

Glenberg, Arthur M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 586-596.

Glenberg, A. M., & Kaschak, M. P. (2002). *Grounding language in action. Psychonomic Bulletin & Review, 9(3), 558-565.*

Goldin, G., & Shteingold, N. (2001). Systems of representations and the development of mathematical concepts. In A. Cuoco (Ed.), *The Roles of Representation in School Mathematics* (pp. 1-23). Reston, VA: National Council of Teachers of Mathematics.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*(3), 1-6.

Hadzigeorgiou, Y., Anastasiou, L., Konsolas, M., & Prevezanou, B. (2009). A study of the effect of preschool children's participation in sensorimotor activities on their understanding of the mechanical equilibrium of a balance beam. *Research in Science Education, 39*, 39-55.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*, 335-346.

Heck, C. & DeFord, B. (2012). Untitled Lesson Study. Unpublished manuscript.

Heefer, A. (2011). Historical objections against the number line. *Science and Education, 20*, 863-880. doi: DOI 10.1007/s11191-011-9349-0

Hefendehl-Hebeker, L. (1991). Negative numbers: Obstacles in their evolution from intuitive to intellectual constructs. *For the learning of mathematics, 11*(1), 26-32.

Hill, H. C., & Shih, T. K. (2009). Examining the quality of statistical mathematics education research. *Journal for Research in Mathematics Education*, 241-250.

Hill, W. H. J. (1968). A physical model for teaching multiplication of integers. *Arithmetic Teacher, 15*, 525-528.

Hurtienne, J., Stößel, C., Sturm, C., Maus, A., Rötting, M., Langdon, P., & Clarkson, J. (2010). Physical gestures for abstract concepts: Inclusive design with primary metaphors. *Interacting with Computers*, *22*(6), 475-484.

International Association for the Evaluation of Educational Achievement (IEA). (2005). *TIMSS* 2003 Assessment. Lynch School of Education, Boston College: TIMSS & PIRLS International Study Center.

Jencks, S. M., & Peck, D. M. (1977). Hot and cold cubes. *Arithmetic Teacher, 24*(1), 71-72.

Johnson, M. (1987). The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago: University of Chicago Press.

Kamii, C., & Joseph, L. L. (1989). Young Children Continue To Reinvent Arithmetic 2nd Grade: Implications of Piaget's Theory. New York: Teachers College Press.

Kamii, C., Lewis, B.A., & Kirkland, L. (2001). Manipulatives: When are they useful? *Journal of Mathematical Behavior, 20*, 21-31.

Kilhamn, C. (2011). Making Sense of Negative Numbers (Vol. 304). Gothenburg: University of Gothenburg.

Kontra, C., Lyons, D.J., Fischer, S.M., & Beilock, S.L. (2015). Physical experience enhances science learning. *Psychological Science, on-line first*, 1-13. doi: 10.1177/0956797615569355

Küchemann, D. (1981). In Hart, K. M., Ed. (1981). Positive and negative numbers. *Children's understanding of mathematics: 11-16*: John Murray London, p. 82-87.

Lakoff, G., & Nunez, R. E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being.* New York: Basic Books.

Lappan, G., Fey, J., Fitzgerald, W., Friel, S., & Phillips, E. D. (2006). *Connected Mathematics 2.* Boston: Pearson- Prentice Hall.

Liebeck, P. (1990). Scores and forfeits-an intuitive model for integer arithmetic. *Educational Studies in Mathematics, 21*, 221-239.

Linchevski, L., & Williams, J. (1999). Using intuition from everyday life in 'filling' the gap in children's extension of their number concept to include the negative numbers. *Educational Studies in Mathematics, 39*(1/3), 131-147.

Math is Fun. *Multiplying negatives*. Retrieved on March 24, 2013 from http://www.mathsisfun.com/multiplying-negatives.html

Moreno, R., & Mayer, R. E. (1999). Multimedia-Supported metaphors for meaning making in mathematics. *Cognition and Instruction, 17*(3), 215-248.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards Mathematics*.

Nunez, R.E, Edwards, L.D., & Matos, F. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics, 39*(1/3), 45-65.

Nunez, R. E. (2000). *Mathematical idea analysis: What embodied cognitive science can say about the human nature of mathematics.* . Paper presented at the Opening plenary address in Proceedings of the 24th International Conference for the Psychology of Mathematics Education. Retrieved from retrieved from nunez' website

Nunez, R.E. (2012). On the science of embodied cognition in the 2010s: Research questions, appropriate reductionism, and testable explanations. *Journal of the Learning Sciences, 21*, 324-336.

Nurnberger-Haag, J. (2007). Integers made easy: Just walk it off! *Mathematics Teaching in the Middle School, 13*(2), 118-121.

Nurnberger-Haag, J. (November, 2013) Metaphors students express about integers after using multiple models. Poster presented at the *International Group of the Psychology of Mathematics Education-North American Chapter Conference*, Chicago, IL.

Nurnberger-Haag, J. (August, 2014). Noticing verbally and physically expressed conceptual metaphors for mathematics: A method for conceptual metaphor analysis. Poster presented at the *9th International Conference On Conceptual Change, Special Interest Group of the European Association of Research on Learning and Instruction*. Bologna, Italy.

Nurnberger-Haag, J. (April 2014). 'Model-Movements' matter: What students learn about number by packing, turning, and hopping. Paper presented in symposium entitled "Theorizing movement and movement-based methods in embodied mathematics learning" chaired by Susan Gerofsky with Discussant, Nathalie Sinclair at the *Annual Meeting of the American Educational Research Association*, Philadelphia, PA.

Periasamy, E., & Zaman, H. B. (Nov 11-13, 2009). *Augmented reality as a remedial paradigm for negative numbers: Content aspect.* Paper presented at the Visual Informatics: Bridging Research and Practice, First International Visual Informatics Conference, IVIC Kuala Lumpur, Malaysia.

Pierson Bishop, J., Lamb, L.L., Philipp, R.A., Whitacre, I., Schappelle, B.P., & Lewis, M.L. (2014). Obstacles and affordances for integer reasoning: An analysis of children's thinking and the history of mathematics. *Journal for Research in Mathematics Education, 45*, 19-61.

Ponce, G. A. (2007). It's all in the cards. *Mathematics Teaching in the Middle School, 13*(1), 10-17.

Roth, W.-M. (2001). Gestures: Their role in teaching and learning. *Review of Educational*

Research, 71(3), 365-392.

Rotman, B. (1993). *Signifying nothing: The semiotics of zero*. Stanford, CA: Stanford University Press.

Rousset. (2010). *Epistemic fidelity of didactical models for the teaching of negative numbers.* Unpublished Master's Thesis?, University of Melbourne, Melbourne.

Ryan, J., & Williams, J. (2007). *Children's Mathematics 4-15: Learning from Errors and Misconceptions*. Berkshire, England: Open University Press.

Sassenberg, U., & van der Meer, E. (2010). Do we really gesture more when it is more difficult? *Cognitive Science, 34*, 643-664.

Scott Foresman-Addison Wesley (2011). enVisionMath Grade 5, Glenview IL: Author.

Semadeni, Z. (1984). A principle of concretization permanence for the formation of arithmetical concepts. *Educational Studies in Mathematics, 15*, 379-395.

Sfard, Anna. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Sfard, A. (2000). Symbolizing mathematical reality into being–or how mathematical discourse and mathematical objects create each other. In Cobb, P., Yackel, E., & McClain, K. (Eds.).*Symbolizing and Communicating in Mathematics Classrooms, Lawrence Erlbaum, Mahwah, New Jersey and London*, 37-98.

Sfard, A. (2007). When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint. *Journal of the Learning Sciences, 16*(4), 565-613.

Shadish, W.R. (2012). Randomized Experiments, Propensity Scores, Regression Discontinuities, and Other Ways of Making Good Causal Inferences in Experiments with Human Participants. *MQM Seminar Series*. November 16, 2012. East Lansing, MI.

Star, J., & Nurnberger-Haag, J. (2011). *Toward a research agenda on mathematical models*. Paper presented at the Annual Meeting of the American Educational Research Association.

Tabachnick, B.G., Fidell, L.S. (2013). Using Multivariate Statistics, sixth edition. US: Pearson.

Stephan, M., & Akyuz, D. (2012). A proposed instructional theory for integer addition and subtraction. *Journal for Research in Mathematics Education, 43*(4), 428-464. doi: 10.5951/jresematheduc.43.4.0428

Tavakol, & Dennick. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55.

Thompson, P. W., & Dreyfus, T. (1988). Integers as transformations. *Journal for Research in Mathematics Education, 19*(2), 115-133.

Umland, K. (2011). *The role of models in mathematics teaching and learning*. Paper presented at the Annual Meeting of the American Educational Research Association.

University of Chicago School Mathematics Project. (2007). *EveryDay Mathematics Grade 5* (3rd ed.). Chicago: McGraw-Hill.

Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. In A. Coxford and A.P. Schulte (Eds.), *The ideas of algebra, K-12* (p. 8-19). Reston, VA: NCTM.

Van de Walle, J., Karp, K., & Bay-Williams, J. (2010). *Elementary and Middle School Mathematics: Teaching Developmentally* (7th ed.). New York: Allyn and Bacon.

Vig, R., Murray, Eileen, & Star, Jon R. (2014). Model breaking points conceptualized. *Educational Psychology Review, 26*, 73-90.

Vlassis, J. (2002). The balance model: Hindrance or support for the solving of linear equations with one unknown. *Educational Studies in Mathematics, 49*(3), 341-359.

Vlassis, J. (2004). Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and Instruction, 14*, 469-484.

Vlassis, J. (2008). The role of mathematical symbols in the development of number conceptualization: The case of the minus sign. *Philosophical Psychology, 21*(4), 555-570.

Warfield, Janet, & Meier, Sherry L. (2007). Student performance in whole-number properties and operations. In P. Kloosterman & F. K. J. Lester (Eds.), *Results and Interpretations of the 2003 Mathematics Assessment of the National Assessment of Educational Progress* (pp. 43-66). Reston, VA: National Council of Teachers of Mathematics.

Yelon, S. L., Ford, J. K., & Golden, S. (2013). Transfer over time: Stories about transfer years after training. *Performance Improvement Quarterly*, 25(4), 43-66.