Changing mental representations using related physical models:
The effects of analyzing number lines on learner internal scale of numerical magnitude

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Barbara J. Bengtson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

William M. Bart

August 2013

Acknowledgements

I would like to thank my adviser, Dr. Bart, for providing opportunities to teach and learn from my fellow graduate students as a graduate assistant and professor for the Education of the Gifted and Talented and Creativity and Intelligence courses. I enrolled in these courses outside of my graduate field; the content and discussions inspired me to reconsider my chosen field and submit an application to the Department of Educational Psychology. Dr. Bart's questions nurtured my problem solving and reflection skills.

My examining committee members also deserve gratitude. Dr. Sashank Varma shared expertise and concise knowledge in psychology, mathematics, problem solving, and article review. His feedback, both challenging and supportive, facilitated persistent progress. Dr. Terry Wyberg's supportive questioning and mathematics education expertise is also greatly appreciated.

I would like to thank my husband Steven Bengtson, for patience and support while I considered and pursued this endeavor, and our children, Haley and William, for providing inspiration and prompting numerous questions in my continuing quest to understand complex learning processes.

Finally, I would like to thank both Dr. Cathy Wick and Mr. Mark Ziebarth for supporting my continued professional growth as an educator and learner. Dr. Wick introduced new ways to explore math; she is an excellent role model and source of professional knowledge. I continue to benefit from Mr. Ziebarth's innovative educational leadership as a principal and fellow instructional strategy enthusiast.

Dedication

This dissertation is dedicated to my friends who continue to learn and share their knowledge.

Abstract

Understanding the linear relationship of numbers is essential for doing practical and abstract mathematics throughout education and everyday life. There is evidence that number line activities increase learners' number sense, improving the linearity of mental number line representations (Siegler & Ramani, 2009). Mental representations of numbers of children in kindergarten through 2nd grade were examined. Methods of improving mental representation using number line activities were also examined. This experimental study included a pretest, interventions, control, and posttest. Analyses were completed to determine accuracy and linearity of estimation patterns as a reflection of mental representations before and after interventions. Age and achievement test data analyses contributed developmental and mathematics performance information to the study.

The findings of this study support existing research indicating children's understanding of number improves with age, with accurate and linear mental representations on a 1 to 100 number line fitting few students in kindergarten, about half in 1st grade and the majority of students in 2nd grade.

This study also contributes to our understanding of the educational power of number line activities and interventions. After three short experiences with board games and broken number line puzzles growth is evident, although not significant, for performance on the number line estimation task. Further study must be done to add to our understanding of number line estimation as well as the activities which improve mental representations.

Table of Contents

List of Tables	viii
List of Figures	ix
Chapter 1: Introduction	**1**
Statement of the Problem	1
Overview of the Study	1
Chapters in the Dissertation	2
Chapter 2: Review of Literature	**2**
Understanding number	2
Early Numeracy and Continuing Achievement	4
Log-to-Lin Shift and Achievement in Mathematics	6
Teaching Number Sense	10
Study Purpose	13
Chapter 3: Method	**15**
Participants	15
Materials	15
Number Line Estimation Task	15
Linear Board Game	17
Broken Number Line Puzzles	19
Achievement Data	21
Design and Procedures	22
Design	22

Procedures 23

Mean percent absolute error 25

Degree of linearity 25

Mathematical achievement 26

Chapter 4: Number Line Estimation Baseline Results **27**

Predictions 27

Grade and Baseline Mean Percent Absolute Error 28

Grade and Baseline Linearity of Estimates 28

Baseline Assessment Variability 29

Discussion 30

Chapter 5: Activities Using Physical Models **34**

Predictions 34

Baseline Assessment 35

Condition, Grade, and Achievement on Mean Percent Absolute Error 35

Condition, Grade, and Achievement on Patterns of Estimates 36

Change from Baseline to Post Assessment 36

Condition, Grade, and Achievement on Mean Percent Absolute Error 36

Condition, Grade, and Achievement on Linearity of Estimates 37

Experimental Condition and Changes in Number Line Estimates 37

Condition and Change in Mean Percent Absolute Error 37

Condition and Change in Linearity of Estimates 37

Grade, experimental condition, and change in number line estimates 38

Kindergarten changes in mean percent absolute error 38

Kindergarten changes in linearity of estimates 38

1st grade changes in mean percent absolute error 39

1st grade changes in linearity of estimates 39

2nd grade changes in mean percent absolute error 40

2nd grade changes in linearity of estimates 40

Linear Board Game Discussion 40

Broken Number Line Puzzle Discussion 42

Changes in Number Line Estimates Controlling for Achievement 43

Low achievement and changes in mean percent absolute error 43

Low achievement and changes in patterns of estimates 44

Average achievement and changes in mean percent absolute error 44

Average achievement and changes in patterns of estimates 45

High achievement and changes in mean percent absolute error 45

High achievement and changes in patterns of estimates 45

Discussion 46

Chapter 6: Discussion: Number Line Activities as Interventions 48

Hypotheses 48

Absolute error rates and linearity vary by grade 49

Changes in understanding from baseline to post assessment 49

Limitations of this study 51

Variability in field collection of data 52

Chapter 7: Conclusions **54**

References **58**

List of Tables

Table 1 *Patterns of Responses for Number Line Estimation Task by Grade* 29

Table 2 *Linear Board Game Descriptive Statistics for Number Line Estimation Task Response Variables by Grade* 41

Table 3 *Number Line Puzzle Descriptive Statistics for Number Line Estimation Task Response Variables by Grade* 42

List of Figures

Figure 1 *Number Line Estimation Task Examples* 16

Figure 2 *Linear Board Game Examples* 18

Figure 3 *Number Line Puzzle Examples* 20

Figure 4 *Baseline Assessment Variability* 30

CHAPTER 1

Introduction

Statement of the Problem

Students with poor understanding of number cannot meet increasingly difficult demands in the mathematics classroom. The National Mathematics Advisory Panel (NMAP) and the National Council of Teachers of Mathematics (NCTM) both address the importance of the mathematical area known as *number*. The NMAP Final Report (2008) boldly states that “poor number sense interferes with learning algorithms and number facts and prevents use of strategies to verify if solutions to problems are reasonable (p27).” NCTM’s Standards document (2000) remarks that number is crucial to all areas of mathematics. Magnitude is an essential feature of number which impacts estimation as well as computation across mathematical contexts. Understanding learner perceptions of magnitude informs instruction; providing improved mental models will increase student ability to verify solutions and grow mathematically.

Overview of the Study

This experimental study contributes to the understanding of mental number line models for students in kindergarten through 2nd grade. It is designed to evaluate activities designed to help learners develop more accurate internal representations. A pre-post assessment design with measures of magnitude representations is used with an experimental group and control group. Magnitude representations are measured using a number line estimation task. Participant activities with whole numbers are in the form of small group games using a number line board game or number line puzzles. The questions this experimental design addresses are: (1) How accurately do students in kindergarten, 1st grade and 2nd grade estimate numbers for locations on a line given only endpoint values? (2) Which model (linear or logarithmic) best fits student's mental representations of the number line? (3) Does utilizing physical models shift students' mental representations of number lines to be more linear and less logarithmic?

Chapters in the Dissertation

The dissertation is comprised of six chapters following this introduction. Chapter 2 is a review of literature which includes research focused on a) the importance of number, b) measuring understanding of the base-ten number system, and c) interventions related to developing number sense. Chapter 3 describes the methodology used in the dissertation including descriptions of participants, materials, as well as design and procedures. Chapter 4 describes the baseline assessment data and the relation to results of existing research. Chapter 5 describes the results of a number line board game and broken number line puzzle interventions and discusses activities using number lines. Chapter 6 revisits hypotheses, the results and describes limitations. Chapter 7 is the conclusion.

CHAPTER 2

Review of Relevant Literature

Understanding Number

The National Mathematics Advisory Panel (2008) describes early number sense as "an ability to immediately identify the numerical value associated with small quantities…a facility with basic counting skills, and a proficiency in approximating the magnitudes of small numbers of objects and simple numerical operations" (p. 27). This description is reflected in mathematics experiences in preschool through primary grades (kindergarten, grade 1, and grade 2) when math is often taught and practiced through counting in series, writing numbers, and counting objects (Wright, Martland, Stafford, & Stanger, 2006). The skills associated with these activities are often rudimentary and isolated, taught only to facilitate basic fact computation using rote memorization and practice; development of deeper understanding of number is essential for future mathematics (Baroody, Bajwa, & Eiland, 2009).

According to the National Council of Teachers of Mathematics (NCTM) Principles and Standards for School Mathematics (2000), prekindergarten through grade 12 instructional programs should help students "understand numbers, ways of representing numbers, relationships among numbers, and number systems" (p. 78). This standard emphasizes understanding and relationships, hallmarks of number sense.

Kastberg and Walker (2008) invited college students preparing to be teachers to "draw a picture of one million of something" as an exam question. Twenty percent of the students drew pictures which only suggested that a million was very large, such as a scene of a beach with labels on a head of hair, bottle of sand, and ocean with fish. Some drew objects one for one, some used slash marks or icons, and others did not attempt to represent quantity. Finally, some used mathematical abstractions, such as repeated addition or icons representing larger numbers. The difficulties with representations may indicate that large numbers are not directly observed quantitatively and are outside of everyday experience; additionally, participants that were successful used a variety of strategies to represent solutions. The representations from adults, who have formal education experiences, demonstrated how numbers are used to make sense of diverse concepts and the complex and nuanced characteristics of numerical representations.

Early Numeracy and Continuing Achievement

Because number is a complex construct, it is useful to organize relatively discrete number skills; one of these organizations is *number competencies*. Described by Jordan, Kaplan, Ramineni, and Locuniak (2009), number competencies depict diverse early math abilities which include counting principles, joining and separating sets, immediate apprehension of values of small quantities (subitizing), number comparisons, and magnitude. Counting principles include (a) understanding that the final number when counting a set is the total number of items in the set and (b) counting order is irrelevant, as long as each item is counted once. Joining and separating sets includes the physical as well as verbal understanding that, for example, 3 and 2 make 5; this is foundational for adding and subtracting.

Number competencies are important for later mathematical development. Jordan et al. (2009) found that measures of children’s number competencies in kindergarten predicted their rate of growth and math achievement in third grade. Early number competencies were measured using a number competency core battery developed for a larger study. The battery included counting and number recognition, number comparisons, nonverbal calculations (problems were presented with a count for a set of chips which were hidden, a given number of chips were either added or subtracted and the child named the sum or difference), oral story problems, and oral addition and subtraction problems. The measure of math achievement in this study was the Woodcock-Johnson Math, Calculation and Applied Problems portions. All correlations between the measures were significant. A sequential growth curve model was determined and regression analyses show three modest, but statistically significant relationships (a) higher average performance on number competence in kindergarten is associated with higher average performance on WJMath in 3rd grade, (b) steeper growth in number competence over kindergarten until the middle of first grade is associated with higher average grade 3 WJMath performance, and (c) higher average performance on number competence in kindergarten is associated with higher rates of growth on WJMath between first and third grade. Approximately 66% of the variance in WJMath performance at the end of third grade and 10% of the variance in the WJMath growth can be accounted for by the number competence growth factors.

Geary (2011) found further evidence for the importance of early numeracy skills to later mathematics achievement in his 5-year longitudinal study. Geary measured understanding of counting principles, addition strategies, number sets, and number line estimation and compared these early numeracy measures to standardized assessments of intelligence, working memory and mathematics (Numerical Operations). He separated the numeracy measures and found (a) children who performed above average on the number sets variable and used efficient counting and decomposition strategies for complex problems at the beginning of 1st grade maintained a consistent advantage on the Numerical Operations test and (b) a number line slope and addition retrieval slope (for simple problems) effect showed benefits which were small early, but increased with successive grades. A score one standard deviation above average on the number line task at the beginning of 1st grade resulted in a non-significant 0.14 point advantage on the Numerical Operations test at the end of 1st grade, but this improved to a 0.79 point advantage by the end of 5th grade. Ability in early numeracy relates to future mathematics achievement effectively demonstrating the foundational quality of number.

Log-to-Lin Shift and Achievement in Mathematics

Existing research utilizing the number line estimation task provides a wealth of information related to developmental understanding of magnitude (i.e., Ashcraft & Moore, 2012; Booth & Siegler, 2006, 2008; Cohen, 2011; Siegler & Booth, 2004). The prompt for a number line estimation task is a horizontal line which is unmarked except for the endpoints and one point or number provided as a target. The target is either a mark present on the line for which the student will supply a corresponding number (Point to Number method; PN) or a number is supplied requiring the student to make a corresponding mark on the line (Number to Point method; NP). Responses to number line estimation prompts describe the participant's mental number line and provide insight to the related understanding of numerical magnitude. Children in kindergarten through 2nd grade completing number line estimation tasks with endpoints of 0-to-100 demonstrated a logarithmic understanding of magnitude that evolved to become a more linear understanding around 2nd grade. This developmental change from logarithmic to linear understanding is often referred to as the *log-to-lin shift*. Additionally, linear understanding of number has been correlated to high math achievement in many studies (i.e., Ashcraft & Moore, 2012; Booth & Siegler, 2006, 2008).

Siegler and Booth (2004) provided evidence of the log-to-lin shift and the relationship between linearity of internal representation and math achievement. The kindergarten, 1st grade and 2nd grade students in that study completed number line estimation tasks on 0-to-100 lines. The estimations of kindergartners were significantly farther from the actual number than were those of 1st or 2nd graders (percent absolute error = 24%, 14%, and 10%, respectively). Regarding the pattern of estimates (a) median estimates of kindergartners were better fit by the logarithmic function than by the linear function (R^2 = 0.89 vs. 0.69), (b) median estimates of 1st graders were equally best fit by the logarithmic function and linear function (R^2 = 0.94 vs. 0.92), and (c) median estimates of 2nd graders were better fit by the linear function than by the logarithmic function (R^2 = 0.97 vs. 0.85). Critically, the function that was best fit for most children varied with age.

These findings were bolstered by Booth and Siegler (2006) in their study with 2nd and 4th graders. The variance accounted for by the linear function increased by grade from 91% for second grade to 98% for fourth grade. There was also a corresponding reduction in fit for the logarithmic model: (a) 2nd grade both models fit similarly (R^2_{lin} =

0.91 vs. $R^2_{\log}$ = 0.88), *whereas* (b) 4th grade estimates better fit the linear function (R^2_{lin} = 0.98 vs. $R^2_{\log}$ = 0.71).

The relation between estimation task performance and math achievement over early development has been the object of prior studies. Estimation accuracy and linearity were related to math achievement in 1st grade and 2nd grade, but did not reach significance for kindergarten (Siegler & Booth, 2004). In all cases, smaller percent absolute error of estimates accompanied higher math achievement test scores. The correlations for absolute error and achievement for 2nd grade was -0.76 and 1st grade was -0.60; kindergarten was in the same direction, but not significant (-0.32). The negative correlation indicates that lower absolute error is related to higher achievement. Similar relationships were found between linearity and achievement; 2nd grade (0.81), 1st grade (0.54), and kindergarten (0.29; *ns*). Booth and Siegler (2006) also found that both low absolute error and high linearity were related to higher achievement in their study with 2nd and 4th grade students, and concluded that student use of logarithmic understanding increased difficulties in computation and estimation of appropriate answers.

More recently Ashcraft and Moore (2012) replicated the number line estimation task using a computer and determined how latencies compared to linearity and error rates. They tested children in 1st to 5th grades and college students and found that latencies decreased with age while replicating evidence of the log-to-lin shift with a decrease in errors and increase in linearity with age. Estimates of 1st graders were an equal fit to logarithmic and linear functions (R^2_{lin} = 0.69 vs. $R^2_{\log}$ = 0.69) and 2nd grader estimates were better fit to the linear function (R^2_{lin} = 0.85 vs. $R^2_{\log}$ = 0.78). Universally, 3rd graders to adult had best fit to the linear function with mean R^2_{lin} values of 0.93, 0.95, 0.96, and 0.98 for 3rd, 4th, 5th grades and college, respectively. In the second part of the study participants in 3rd grade, 4th grade, 5th grade, and college completed another number line estimation task with endpoints of 0 and 1000. The change to the larger magnitudes resulted in some 3rd, 4th, and 5th graders, who were better fit to the linear function in the 0-to-100 task, appearing to regress with 0 to 1000 estimates which better fit the logarithmic function (20%, 26%, and 17%, respectively); adults were still all best fit with the linear function. Ashcraft and Moore concluded that age related changes in number line estimations are influenced by magnitude understanding.

Rather than logarithmic, another possibility for naïve respondent mental number lines is a multi-linear function. Ebersbach, Luwel, Frick, Onghena, and Verschaffel (2008) provide evidence for a 2-segment model with the change point occurring where children lack familiarity with the numbers and errors explained by the accumulator model which posits that absolute error increases with magnitude. Moeller, Pixner, Kaufmann, and Nuerk (2009) also prefer the 2-segment linear model, but argue for a change point related to transitions near decades and that increased understanding of linearity is due to improved integration of units and tens. Multi-linear models are refuted by Ashcraft and Moore (2012) who described an M-shaped graph of absolute errors for more linear thinkers, indicating midpoint strategy use. Knowledge of the midpoint strategy is evidenced by the M-shaped graph of absolute error; more accurate estimates result in error values closer to zero near the endpoints and center of the number line while errors increase as distance from these three benchmarks increases. By contrast, less knowledgeable students have a tent-shaped error graph, with midpoint estimates farthest from the actual number using only endpoints for reference. This interpretation of response patterns provides insight for instructional opportunities with potential to shift learner mental number lines to more linear forms.

Teaching Number Sense

Of the number competencies, research provides ample evidence that understanding numerical magnitude is particularly useful to continued mathematical learning (Jordan et al., 2009). Number lines are primarily a representation of magnitude but take many forms for many purposes in mathematics education. At the elementary level number lines are often used to represent whole number magnitudes, but have been adapted to represent addition and subtraction as well as for determining elapsed time, a difficult task for young children (Dixon, 2008). Bay (2001) provides multiple uses for empty number lines made of rope as an interactive visual representation to help students understand relative magnitudes of large numbers (0-1,000,000 number line) and fractions (0-1 number line), and to solve algebraic problems. In high school number line activities are used in constructions to understand magnitudes of irrational numbers (Coffey, 2001). Izsák, Tillema, and Tunç-Pekkan (2008) provide evidence that number lines support both the teaching and learning of fraction addition for knowledgeable students, but discover that points on a line partitioned to represent fractions are difficult for less knowledgeable students to process. As evidenced by the longitudinal data provided by Geary (2011) research has shown that less knowledgeable students tend to remain in that group;

providing appropriate support so that all students can grow is the goal of representation focused teaching strategies.

Studies have demonstrated that learners struggling early in math continue to have low growth in mathematics achievement and that a naïve understanding of numbers relates to income differences (Geary, 2011; Ramani & Siegler, 2011; Seigler & Ramani, 2009). Jordan et al. (2009) found that measures of children's number competencies in kindergarten predicted their rate of growth and math achievement level in third grade and 20% of the variation in achievement and 7% of the variance in growth were accounted for by gender, income and kindergarten start age. These differences are present in most educational settings. Number line estimation provides an excellent resource for measuring effects of similarly structured instructional activities designed to overcome mathematical achievement deficits related to demographic differences.

Siegler and Ramani conducted studies to compare circular and linear number line game boards for number sense development. A subsequent study further explored the use of the board games while considering the needs of students in different types of programs. The first study was comprised of students in Head Start, a preschool program which serves very low income families, and nearby neighborhood childcare centers (Siegler & Ramani, 2009). The follow up study included students from six predominantly middle- or upper-middle-class preschools, three of which had university affiliations (Ramani & Siegler, 2011). The samples in these two studies consisted of participants with families of different socio-economic backgrounds. Another difference was the age of children in the preschool programs; the low-income programs included children from 4 years 0 months to 5 years 5 months, whereas the middle-income group included an age range from 3 years 5 months to 4 years 8 months.

Results indicated that the linear version was significantly more effective for improving magnitude understanding and increased learner ability to benefit from numerical activities. Initial number line estimation analyses indicated that the children from middle-income homes came with greater numerical knowledge. Though these students averaged eight months younger than the students from low-income homes, their estimates were more linear (mean $R^2_{\text{lin}} = 0.27$ versus 0.15). Matching was done to allow for comparison and allowed for identification of students whose pretest performance did not differ on any of the tasks. Linearity increased for both of these groups after playing the linear board game. For the children from low-income families linearity increased from 16% on the pretest to 38% on the posttest; for the children from middle-income families linearity increased from 17% on the pretest to 29% on the posttest. The effect sizes demonstrate a large effect on the knowledge of the low-income group compared to a medium effect on that of the middle-income group. Also notable was variation based on level of initial knowledge; linearity increased from pretest to posttest for those below the median on the pretest, but not for those who scored above the median. The lasting effect of improved ability to estimate magnitude is moving to a linear understanding of number; this provides a method for educators to reduce the achievement gap which exists between low- and middle- income children (Ramani & Siegler, 2011). It is important to note that the use of a game with similar attributes to the mental model was most effective for improving the mental model.

Ashcraft and Moore (2012) identify two processes that have immediate impact on participant performance on the number line assessment (a) instruction specific to the number line characteristic of even intervals and (b) the midpoint identification strategy. Even intervals instruction was implemented by Siegler and Ramani (2009) and is described by Ashcraft and Moore (2012) as minimal teaching with a profound effect. Additionally, the midpoint strategy, which produces the M-shaped graph of absolute errors, was used as an orienting trial with feedback to increase accuracy and linearity of children's estimates (Siegler & Booth, 2004).

Study Purpose

Magnitude is an essential feature of number which impacts estimation as well as computation across mathematical contexts. It is important to improve mental representations of magnitude to enhance students' capacity to check work and grow mathematically. This dissertation investigates two research questions concerning children's mental representations of number magnitude: (1) How accurately do students in kindergarten, 1st grade and 2nd grade estimate numbers for locations on a line given only endpoint values? (2) Which model (linear or logarithmic) best fits student's mental representations of the number line? Answering these questions provide a baseline and insight to potential student strengths and needs for understanding magnitude.

The goal of education is to support student growth. Once understanding is established, methods for teaching and learning must be implemented and evaluated. Therefore, the third – and primary research question concerns intervention: (3) Does utilizing physical models shift students' mental representations of number lines to be more linear and less logarithmic?

CHAPTER 3

Method

Participants

Data were collected from a convenience sample of 95 students in Kindergarten, 1st grade, and 2nd grade, aged 5-8 years-old. This age range was selected because there is evidence that a developmental shift from logarithmic to linear mental number lines for the number range 0-to-100 occurs in 2nd grade (Siegler & Booth, 2004). Participants included 22 children in kindergarten (mean age = 6.1 years, *SD* = 0.2 years), 31 in 1st grade (mean age = 7.0 years, *SD* = 0.2 years), and 42 2nd grade students (mean age = 8.0 years, *SD* = 0.3 years). The children attended an elementary school in a small city (population approximately 5200) in a midsize school district serving approximately 5000 students. Sixteen students are not included as participants because they were absent from the baseline administration of the number line estimation task (11 students: 3 kindergarten, 4 first grade, and 4 second grade) or they missed one or more items when completing the baseline or post number line estimation task (5 students: 1 kindergarten, 2 first grade, 2 second grade). There were 5 cooperating teachers. For contributing instructional time to the study, educators were debriefed on results and provided intervention protocols and materials for future use. Students received no extrinsic compensation.

Materials

Number line estimation (NLE) task. A 25 cm number line was displayed with 0 and 100 marked on opposite ends with vertical hatch marks, each labeled slightly below the line; a target number appears on the top right of the page (see Figure 1). There were 26 trials of number line estimation using the target values: 3, 4, 6, 8, 12, 14, 17, 18, 21, 24, 25, 29, 33, 39, 42, 48, 52, 57, 61, 64, 72, 79, 81, 84, 90, and 96. These values are identical to those used by Ashcraft and Moore (2012) and Siegler and Booth (2004).

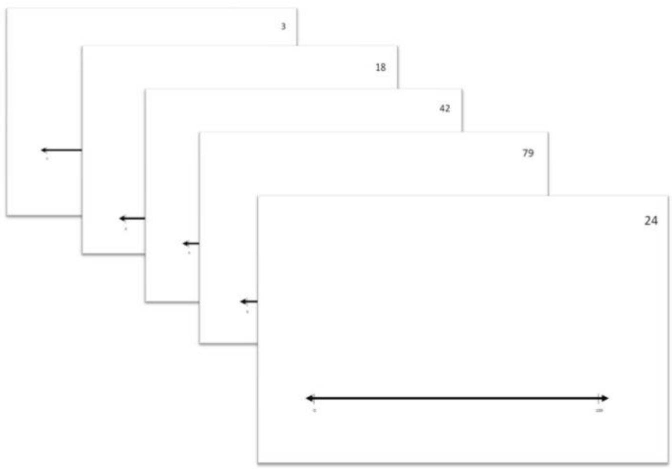

Figure 1. Number line estimation task samples. Each page consisted of a number prompt on the top right and a line marked with 0 on the left and 100 on the right. The task included 26 different pages stapled in the top left corner.

The estimation task was in the form of a packet of sheets printed landscape and stapled on the top left; the target values were presented to all participants in 1 of 3 randomized orders before the first and after the third intervention sessions. Two demonstration trials, using 0 and 100 as target numbers, preceded the administration of the task to ensure that students could effectively make a clean, single mark on the line. On each page students placed a mark across the horizontal line to indicate where the value belongs. The experimenter scored each page using a transparent overlay of the number line with 101 hatch marks numbered at zero and each decade to 100. Student marks were assigned the number corresponding to the nearest hatch mark on the overlay; those between overlay hatch marks were assigned the higher value of the two. The mark placements made for each value on this instrument demonstrated the accuracy (absolute error in estimates) and degree of linearity (patterns of the series of estimates graphed against the expected value) of the participant's mental number line.

Baseline and post-test measures consisted of the number line estimation procedures described above. Assessments provide evidence to describe participant whole number magnitude understanding from 0 to 100. Studies have shown the number line estimation method of assessment to correlate to other measures of student achievement in math.

Linear board game. This game was adapted from those described by Siegler and Ramani (2009, 2011) and Whyte and Bull (2008) and featured a game board with sequentially numbered spaces in a linear format and a number cube with equal likelihood of rolling one or two (see Figure 2). The game board was available in four leveled versions for each session (*a*, *b*, and *c*). The first level of each linear board game included a sequence with a decade ending: (*a*) 1 to 10, (*b*) 41 to 50, and (*c*) by tens from 0 to 90 for the first, second, and third intervention days, respectively. The second level for each session was the same sequence used in the first level with numbers missing. The second level missing numbers were (*a*) 2, 5, 6, and 9; (*b*) 42, 46, and 48; and (*c*) 20, 60, and 80. The third level of each session was a sequence starting in the same decade as the previous sequence but going into the next decade: (*a*) 6 to 15, (*b*) 48 to 57, and (*c*) by tens from 6 to 96. The final activity available for each session used the previous sequence with numbers missing. The fourth level missing numbers were (*a*) 7, 11, 13, and 14; (*b*) 49, 52, and 56; (*c*) 16, 56, 76, and 86.

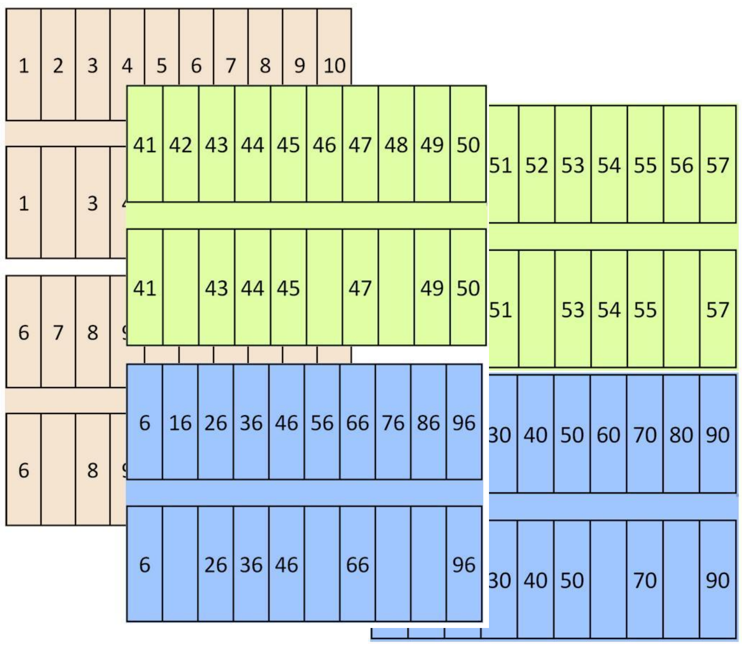

Figure 2. Linear board game samples. Each game board consisted of ten spaces with four board variations color-coded by session.

Prior to playing protocol dictated that students select their playing piece and read the numbers in chorus to ensure all players were familiar with the sequence. The board game was turn based with a progression to the finish and was played multiple times with students taking turns being first (the advantage goes to the first player). The student whose first name is alphabetically first began by rolling the die then moved the selected piece. After rolling the die, the player counted aloud, reading each number as it was passed. Any mistakes were corrected and the count redone. Play continued as the three to four students in the group took turns rolling the die and moving, stating the numbers they passed; the first student to the end of the board won the round. After completing a round the student sitting to the left of the starting student began a new round, because there is an advantage to going first. After one or two errorless rounds on each board, the board was changed; the second board had the same sequence as the first with numbers missing. The second and fourth versions have missing numbers and were played in the same manner as the complete sequences; students mentally filled in the missing numbers and stated the number that belonged in the unlabeled space. Play continued for ten minutes, changing game boards with sequences of increasing magnitude as mastery was demonstrated.

The first version, numbered one to ten with no missing digits, has been proven successful at increasing linearity of mental number line representations in the studies by Siegler and Ramani (2009, 2011), as well as Whyte and Bull (2008).

Broken number line puzzles. This task was designed by the researcher to increase physical and cognitive interactions with number lines in an effort to support student generation of linear mental representations. Each of the twelve puzzles was a number line with eleven number places marked with vertical hatch marks (see Figure 3). As with the linear board game, four levels were available for each of three sessions. The first level for each session had a decade ending: (a) consecutively numbered from 0 to 10, (b) 40 to 50, and (c) counted by tens from 0 to 100 for the first, second, and third intervention days, respectively. The second level for each day was the same sequence with numbers missing. For each session the missing numbers were (a) 2, 5, 6, and 9; (b) 41, 44, 46, and 48; and (c) 20, 60, and 90. The third level of each session was a sequence starting in the same decade as the previous sequence, but going into the next decade: (a) 6 to 16, (b) 48 to 58, and (c) adding tens from 6 to 106. The final level used level three sequences with missing numbers. The missing numbers are (a) 7, 11, 13, and 14; (b) 49, 52, and 56; and (c) 16, 56, 76, and 86.

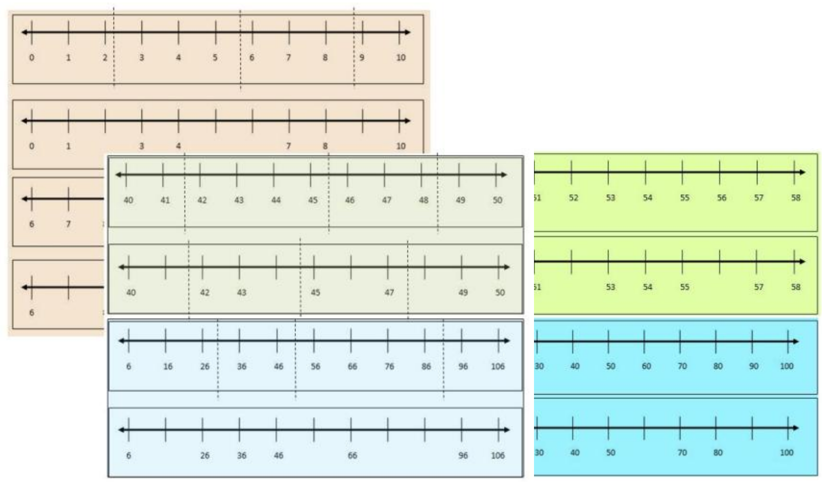

Figure 3. Broken number line samples. Each puzzle consisted of an eleven place number line cut in 4 pieces of varying size. Available each day were four board variations color-coded by session.

This intervention activity featured a number line with 11 hatch marks numbered sequentially and cut into 4 pieces with two to three numbers on each piece. Before building the puzzles the protocol dictated students were shown the complete number line and instructed to read the numbers in chorus to ensure all players were familiar with the sequence. The students were instructed to build the puzzle and when completed, to repeat the sequence silently to self to check work. The number line puzzles were completed simultaneously with comparisons to the solution taking place once all students had solutions; there were multiple versions at each level to allow repeated exposure. In the puzzle activity students were instructed to either silently say the numbers on the puzzle or attempt to identify the middle number of the number line while waiting for all peers to complete the puzzle. Students ordered the pieces to make the number line which was compared with one that had not been separated. Errors were corrected. There were four versions of each broken number line puzzle to enable corrected practice by building the same number line with slightly different break points. Students traded puzzles and built again until there were no errors within the group. Participants first built a puzzle with all of the numbers labeled; its pair had three or four places which have hatch marks but no numbers. Students played in the same manner by mentally filling in the missing numbers and stating the number that belonged in the unlabeled space. Puzzle construction continued for ten minutes, increasing number line magnitude as mastery was demonstrated.

The invention of this activity by the author was a response to a need for a method to improve students' understanding of numbers. Increasing student interaction with and problem solving using number lines provides experience which improves linearity of mental representations. It is included in this study with the linear board game because both have essential characteristics similar to a number line: equally spaced sequential numbers. It was hypothesized that this activity would improve mental number line estimation because it is a physical representation of the number line paired with the cognitive activity required to check accuracy during construction (Ashcraft & Moore, 2012; Marzano, R. J., Pickering, D. J., & Pollock, J. E., 2001).

Achievement data. Grade-appropriate fall math benchmark data were obtained for each participant. Students in kindergarten had scores for the AIMSweb Test of Early

Numeracy Assessment (TEN) which includes subtests for oral counting (OCM), number identification (NIM), quantity discrimination (QDM), and missing number (MNM). Students in first grade had scores for the AIMSweb Mathematics Computation Assessment (M-COMP) which includes items requiring knowledge in column addition, basic facts, and complex computation. Students in second grade had scores for both AIMSweb and NWEA assessments. The AIMSweb Mathematics Concepts & Applications Assessment (M-CAP) provides an overall score comprised of number sense, operations, patterns and relationships, measurement, geometry, and data and probability strands. The NWEA MPG provides an overall composite score (RIT) and sub scores for problem solving, number sense, computation, measurement and geometry, statistics and probability, and algebra.

Design and Procedures

Design. Independent variables were between-subjects and included grade level (k, 1, 2), group condition (linear board game, board game control, broken number line puzzles, or puzzle control), standardized test achievement (above, at, or below grade mean), and baseline number estimation task achievement (above or below grade median). Within-subjects variables included baseline mean percent absolute error and post assessment mean percent absolute error as well as baseline linearity and post assessment linearity analyzed as changes in mean percent absolute error and linearity from the baseline, the dependent variables.

Intervention and control were randomly assigned to classrooms to determine which students will receive the interventions and which students will provide information for experimental controls. 1st grade and 2nd grade each consisted of two classrooms. Students in one classroom at each grade level participated in either the number line game or puzzle condition. The second classroom at each grade level functioned as the control group. Kindergarten had one teacher with three groups. Kindergarten students from each group were randomly assigned to a condition. Students in classrooms that were selected as controls completed the number line estimation task at the beginning and end of the study. The students in the control condition received regular classroom instruction, and no additional structured activity; the board game and puzzle conditions in the study were interventions (additional mathematical experiences) to be compared to performance of students receiving only regular mathematics instruction in classrooms.

Procedure. Participants were grouped based on classroom assignment prior to the baseline number estimation task. All students first completed the number line estimation task described above; these results provide the baseline measure for this study. The number line estimation task was given to students in their regular classroom environment. Students seated at desks or tables received one of three versions of the packet. Prior to distribution packets were ordered so that consecutively distributed tasks were not identical; this was an effort to prevent students from using the work of others to inform their estimates. The administration protocol began with an introduction, students were then instructed to write their name on the top page and wait to ensure everyone was ready for instructions. Each class was told they would look at the number in the top right corner, think about where it belongs on the line below and make a mark across the line to indicate its place. Two demonstrations were completed using sheets with targets of 0 and

100. The number lines in the demonstrations were not labeled. The first target number was zero and was marked near the arrow at the left end of the line while stating, “I think zero belongs here.” The second target was 100 and was marked on a separate sheet on the far right of the number line near the arrow while stating, “I think 100 belongs here.” Students were told that all of the pages they have were labeled with zero and 100 and they would have different numbers to consider and place on the number line; they will look at the number, think about where it belongs, and make a mark across the line in that place. Students continued to do this until all pages in the packet were marked. Errors in procedure, such as students making multiple marks on one page, were immediately corrected. The number line estimation tasks were collected before the experimenter left the room. This process was repeated in each of the classrooms which participated in the study.

Once the baseline was complete, linear board game and broken number line puzzle conditions were randomly assigned to students in classrooms selected to receive interventions; once conditions were assigned, game subgroups were established. Students met in groups of 3 or 4 and played the assigned game or constructed a puzzle for ten minute sessions on three different days, as described above. Each activity level and session moved students through a progression of number lines with increasing magnitudes. Each session included the potential to work on four levels of the group activity. The mastery aspect of the tasks caused variation in the number of completed levels by each group; progress was driven by the group’s ability to successfully complete the assigned activity.

Mean percent absolute error. Each estimate provided was subtracted from the expected value; the absolute value of this difference was the absolute error. Identical to the measure developed by Siegler and Booth (2004), the percent absolute error for each value was calculated by dividing the absolute error by 100 (the scale of estimates). For example, given the target number 39 and a marked estimate corresponding to the actual location of 48, the percent absolute error calculation is $|39-48|/100 = 0.09$; the percent absolute error is 9%. The mean of the percent absolute error was calculated for individuals across estimates and for each value across individuals for between and within group comparisons in analyses.

Number line degree of linearity. Curve fit estimation was done to determine linearity, the expected value functioned as the independent variable, the dependent variable was the corresponding estimate from baseline or post tests. The test estimates were compared against both linear and logarithmic models (accurate estimations fall on the line $y = x$). R^2 was calculated for each model; higher value R^2 indicated that more of the variation in responses was explained by the corresponding model (linear or logarithmic curve formula). The R^2 associated with the linear equation represents the degree of linearity for individual and group estimates. The degree of linearity was compared with the R^2 associated with the logarithmic model, the greater of the two was considered the functional mental model. Using analysis methods replicating Siegler and Booth (2004), each grade's median response for each estimated value was used as the grade's estimated series in linearity analyses. The descriptive data, including mean percent absolute error and linearity, for each cell relating to grade and condition are reported in tables in the results chapter.

Mathematical achievement. Student achievement is variable within grades as well as between classrooms. For this study, student standardized achievement scores were separated into three ordinal categories referring to scores which were low, average (the largest group), and high relative to average scores of students in the same grade; this designation is used as a factor in analyses, as are grade and experimental condition.

CHAPTER 4

Number Line Estimation Baseline Results

Predictions

This study investigated three research questions two of which concern children's mental representations of number magnitude: (1) How accurately do students in kindergarten, 1st grade and 2nd grade estimate numbers for locations on a line given only endpoint values? (2) Which model (linear or logarithmic) best fits student's mental representations of the number line? The third – and primary research question concerns intervention and is addressed in Chapter 5: (3) Does utilizing physical models shift students' mental representations of number lines to be more linear and less logarithmic? Answers to these questions provide a baseline and insight to potential student strengths and needs for understanding magnitude.

Based on research in number magnitude using number line estimation, the following hypotheses were determined in response to questions one and two. Absolute error rates and linearity vary by grade level. Absolute error rates will remain consistent over a four week period in the absence of direct intervention (i.e., control groups), whereas absolute error rates will decrease following three 10-minute number line activity interventions. Linearity of mental number lines will remain consistent over a four week period in the absence of direct intervention (i.e., control groups). Linearity will increase following three 10-minute number line activity interventions; furthermore, dependence on a logarithmic model will decrease.

Grade and baseline mean percent absolute error. The one-way ANOVA, $F(2, 92) = 47.25$, $MSE = 0.01$, $p < .001$, $\eta^2 = 0.51$, indicated differences in mean percent absolute errors across the three grades. The kindergarten mean percent absolute error and related standard error (37%, $SE = 0.033$) were higher than both grade 1 (17%, $SE = 0.019$) and grade 2 (9%, $SE = 0.012$). Mean difference comparisons were significant when comparing 1st grade and 2nd grade to kindergarten at the level $p < .001$($SE = 0.03$ and 0.02, 1st and 2nd, respectively). The 1st to 2nd grade comparison resulted in $p = .002$ ($SE = 0.22$). This is comparable to the 2004 study findings of Siegler and Booth which described the positive relationship between grade and mean percent absolute error; that study found percent absolute error means of 24%, 14%, and 10%, for kindergarten, grade 1, and grade 2, respectively.

Grade and baseline linearity of estimates. Curve estimation analyses were completed for each grade's median estimates for both linear and logarithmic models. A one-way ANOVA was used to determine differences between the linearity measures by grade, $F(2, 92) = 29.79$, $MSE = 0.09$, $p < .001$, $\eta^2 = 0.39$. This indicated considerable grade differences representing the linearity of the number set; degree of linearity tended to increase as grade increased. The kindergarten median estimates corresponded poorly, but had a higher proportion of variability explained with better fit using the logarithmic model, $R_{lin}^2 = 0.47$, $R_{log}^2 = 0.62$. The first grade median estimates fit both models similarly, $R_{lin}^2 = 0.95$, $R_{log}^2 = 0.92$. The second grade median estimates were best fit by the linear model, $R_{lin}^2 = 0.99$, $R_{log}^2 = 0.81$. These outcomes correspond relatively well to the findings of Siegler and Booth (2004): kindergarten, $R_{lin}^2 = 0.69$, $R_{log}^2 = 0.89$, first grade $R_{lin}^2 = 0.92$, $R_{log}^2 = 0.94$, and second grade $R_{lin}^2 = 0.97$, $R_{log}^2 = 0.85$, although kindergarten results provided a notable difference between this study and the 2004 study results.

Additionally, comparisons of linear and logarithmic model fits by individuals at each grade level provide frequencies which support the estimation accuracy trends outlined by the median estimates, depicted in Table 1. Kindergarten students have more estimates which best fit the logarithmic model, first grade students increased the percentage of estimates for which the linear model best fit, and second graders' responses were overwhelmingly best fit by the linear model.

Table 1
Patterns of Responses for Number Line Estimation Task by Grade

	Pattern of Estimates	
Grade	Lin	Log
Kindergarten	45.5	54.5
First Grade	55.0	42.0
Second Grade	93.0	7.0

Note. Lin = percent of response sets with variability best explained by linear model, Log = percent of response sets with variability best explained by logarithmic model

Baseline assessment variability. Pair-wise correlations across all participants among the mean percent absolute error, linearity, and achievement variables indicate significant negative relationships between baseline measures of linearity and mean percent absolute error, r = -.905, p < .001, and achievement level and baseline mean percent absolute error, r = -.223, p = .03. There is a positive correlation between baseline linearity and achievement, r = .343, p = .001. Lower mean percent absolute error values are preferable whereas higher values indicate superior achievement and linearity. The preferred response is clearly illustrated in the top right cell of Figure 4, which shows the location on the graph where high ability 2nd grade students responded with highly linear estimates with low absolute error.

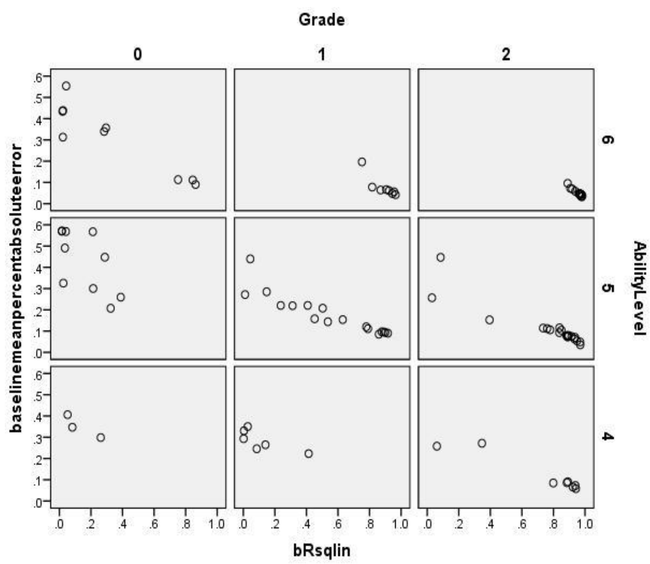

Figure 4 *Baseline Assessment Variability*

Figure 4 as a whole appears to depict a reduction in variability with increased age and achievement (ability level), although regression across all participants using mean percent absolute error and linearity to predict achievement was not robust, $R^2 = 0.16$, *SE* = .648.

Discussion

Baseline results address the research questions: (1) How accurately do students in kindergarten, grade 1 and grade 2 estimate numbers for locations on a line given only endpoint values? (2) Which model (linear or logarithmic) best fits student's mental representations of the number line? The evidence of increased linear understanding by age is related to the decreased mean percent absolute error findings by age. Together these indicators provide evidence of the log-to-lin shift occurring over the primary grades, similar to results of Siegler and Booth (2004).

The protocol for the second experiment number line estimation task in the 2004 Siegler and Booth study included participants locating a point on the number line and receiving feedback from a researcher on an orienting value as an introduction to the estimation task. One design difference in the number line estimation task in this study was prompted by the possible effects of this orienting work and by the research of Ashcraft and Moore (2012) which describes Barth and Paladino's 2011 study in which midpoints were used to both orient participants and explain proportions of the number line. Ashcraft and Moore (2012) provided evidence that even brief number line instruction results in change in performance on the estimation task.

The protocol in this dissertation study oriented students to the activity using the endpoints, which were labeled in the task. This procedural change likely contributed to differences in the kindergarten results when compared to Siegler and Booth (2004). Students made estimates using the endpoints and their personal mental representations; no examples or feedback were given regarding the possible location of any number between 0 and 100. The use of the endpoints as orienting marks was done to prevent new learning at the time of assessment. Students given a target sample between 1 and 99 could potentially use this as a learning opportunity, changing magnitude understanding and the mental representation being assessed. For students with limited exposure to number lines this was a distinct disadvantage and is reflected in the apparent randomness of the estimates. These students placed most estimates on or near an endpoint, rather than near the location of the value. However, the lack of instruction regarding points between zero and one hundred provided an assessment which authentically reflects students' existing mental representations.

The patterns of responses from the naïve students were variable, but effectively demonstrate the results of Ebersbach et al (2008) in regards to end-anchor preferences when individual student response sets are analyzed. When placing a mark to represent the value estimate, four kindergarten response sets were marked near or at 100 for all target values. Five students appeared to have randomly marked either end of the number line. The sole first grader in this group alternated marking 0 and 100 on each page of the task. The orienting activity likely contributed to some students marking estimates at the endpoints. These responses, however, demonstrated varying levels of understanding, which may have been lost if a midpoint location was demonstrated.

The following two baseline responses provide additional insight. Student A exhibited high level end-anchor responses to the estimation task while also inserting two midpoint estimates; when prompted with the value 6, a hash mark was made at the location of 54; when 90 was the value, it yielded a response mark at 51. To further illustrate, sixteen marks were made at the locations of one to four, eight marks were made at the locations of 95 to 99. Most students using the endpoint anchors made their marks on or partially on the 0 and 100 labels; this student made no marks on the endpoints.

Student B responded to all single digit numbers with an estimate of zero and all double digit numbers with an estimate of 100. Contrasting the results provided by these students one can see differences regarding understanding of both magnitude and endpoints. Moeller et al (2009) and Ebersbach et al (2008) proposed a split linear model based on children's familiarity with numbers; the slope of the line associated with familiar numbers would be close to one, the unfamiliar numbers would provide a linear relationship, but with a higher slope.

The number line estimation task can empirically explore the model of the number mental number line held by naïve subjects only if instruction regarding the physical number line is a precursor. The novelty of the number line experience created low validity in kindergarten linearity analyses; however it may illuminate developmental trajectories in numeracy. Overall kindergarten linearity and mean absolute error variability reflect students' understanding experiencing a shift from complete naivety to emergent understanding of the number line at this age level. Grade one and grade two had increasingly linear results which demonstrated that increased exposure to number lines in subsequent years of education eliminated much of the variability evident in kindergarten, the first year in the public school system.

CHAPTER 5

Activities Using Physical Models

Predictions

The third – and primary research question this study addressed concerns number line activities: (3) Does utilizing physical models shift students' mental representations of number lines to be more linear and less logarithmic?

Using linear board games has changed learners' understanding of numbers, resulting in increased linearity of number line estimates (Siegler & Ramani, 2009, 2011). This study utilized similar procedures and game boards in an attempt to replicate these results. Additionally, this research began to investigate the hypothesis that dynamic number line activities provide improved understanding compared to more passive representations. Meta research on strategy instruction for generating and testing hypotheses reveals average effect sizes of .61 and percentile gains of 23 for achievement tests; when learners predict a solution and receive feedback on their prediction, academic achievement improves (Marzano et al, 2001). Puzzles and missing numbers on a number line elicit hypotheses for completion; reciting the numbers on a completed puzzle or number string checks accuracy of puzzle construction and missing number hypotheses.

My final hypotheses for this study are regarding differences in results for the intervention activities. Because the puzzles provide multiple feedback cues, number order as well as the presence of arrows at each end of the puzzle, puzzles will yield stronger results than the board games. Additionally, missing number opportunities will yield stronger results than activities in which all values are given because hypotheses are made and tested.

Baseline Assessment

Both mean percent absolute error and linearity contribute to our understanding of students' mental number line. Analyses were done to determine if random assignment created experimental groups with similarly linear understanding at the time of the baseline assessment. Tables describing number of participants for each experimental condition and grade subgroup and baseline and posttest results are found in the discussion following data from the linear board game and broken number line puzzle activities.

Effect of experimental condition, grade, and mathematics achievement on baseline mean percent absolute error. The ANOVA indicated insignificant interactions between the effects of the experimental condition, grade, and achievement in mean percent absolute errors on the baseline assessment. As described in Chapter 4 there was a main effect of grade which is verified by the ANOVA, $F(2, 60) = 49.829$, $MSE = 0.008$, $p < .001$, $\eta^2 = 0.391$, also present is a main effect of achievement, $F(2, 60) = 9.813$, $MSE = 0.008$, $p < .001$, $\eta^2 = 0.077$. The main effect of experimental condition on baseline mean percent absolute error was not significant. Furthermore, contrasting baseline performance of the linear board game intervention group and the related control group provided similar results, $t(91) = -0.65$, $p = .52$, $d = -0.03$ as did contrasting baseline performance of the broken number line puzzle intervention group and the related control group, $t(91) = -0.29$, $p = .78$, $d = -0.01$. This evidence demonstrates that the different experimental groups started the study with similar estimation skills.

Effect of experimental condition, grade, and mathematics achievement on baseline linearity of estimates. An ANOVA indicated one significant interaction between the effects of the experimental condition, grade, and achievement in linearity on the baseline assessment, that of achievement and grade. The ANOVA, $F(4, 60) = 2.726$, $MSE = 0.066$, $p = .037$, $\eta^2 = 0.055$, also present are the main effects associated with the interaction, the effect of grade, $F(2, 60) = 32.628$, $MSE = 0.066$, $p < .001$, $\eta^2 = 0.331$, and ability, $F(2, 60) = 12.448$, $MSE = 0.066$, $p < .001$, $\eta^2 = 0.126$. The main effect of experimental condition on baseline mean percent absolute error was not significant. Furthermore, contrasting baseline performance of the linear board game intervention group, and the related control group provided similar results, $t(91) = 0.18$, $p = .86$, $d = 0.02$ as did contrasting baseline performance of the broken number line puzzle intervention group and the related control group, $t(91) = 0.08$, $p = .93$, $d = 0.01$. This evidence demonstrates that the different groups started the study expressing similarly linear mental representations.

Change from Baseline Assessment to Post Assessment

Effect of experimental condition, grade, and mathematics achievement on changes in mean percent absolute error. The ANOVA indicated insignificant interactions between the effects of the experimental condition, grade, and achievement in change of mean percent absolute errors. The main effects were also insignificant. The factors in the ANOVA with greatest significance included the interaction between experimental condition and achievement, $F(6, 60) = 1.526$, $MSE = 0.005$, $p = .185$, $\eta^2 = 0.098$, with a slightly larger p-value, the main effect of achievement, $F(2, 60) = 1.981$, *MSE* = 0.005, *p* = .147, η^2 = 0.042, and experimental condition, *F*(3, 60) = 0.912, *MSE* = 0.005, *p* = .441, η^2 = 0.030.

Effect of experimental condition, grade, and mathematics achievement on changes in linearity of estimates. The ANOVA indicated insignificant interactions between the effects of the experimental condition, grade, and achievement in change patterns of estimates. The main effects were also insignificant. The factor in the two-way ANOVA which approached significance was the main effect of grade, *F*(2, 60) = 2.734, *MSE* = 0.030, *p* = .073, η^2 = 0.054.

Experimental Condition and Changes in Number Line Estimates

Experimental condition and change in mean percent absolute errors. The one-way ANOVA, *F*(3, 91) = 1.437, *MSE* = 0.01, *p* = .237, η^2 = 0.04, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that different groups exhibited similar changes in estimation skills. Contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, *t*(91) = 0.06, *p* = .955, *d* < 0.01. However, contrasting performance of the linear board game intervention group and the control group provides somewhat different results, *t*(91) = -1.92, *p* = .06, *d* = -0.04; more significant differences are uncovered when limiting the comparison to these groups.

Experimental condition and change in linearity of estimates. The one-way ANOVA, *F*(3, 91) = 1.331, *MSE* = 0.03, *p* = .269, η^2 = 0.04, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that different groups exhibited similar changes in estimation patterns. Contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(91) = 0.60$, $p = .553$, $d = 0.03$. However, contrasting performance of the linear board game intervention group and the control group provides somewhat different results, $t(91) = -1.86$, $p = .07$, $d = -0.10$; differences appear more significant when limiting the comparison to these groups.

Grade, experimental condition, and changes in number line estimation

Kindergarten experimental condition and change in mean percent absolute errors. The one-way ANOVA, $F(3, 18) = 0.643$, $MSE = 0.01$, $p = .597$, $\eta^2 = 0.10$, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that kindergarten groups exhibited similar changes in estimation skills. Furthermore, contrasting performance of the linear board game intervention group and the control group provided similar results, $t(18) = -0.545$, $p = .593$, $d = -0.03$. Contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(18) = -0.946$, $p = .357$, $d = -0.53$, although significance of difference increased slightly.

Kindergarten experimental condition and change in linearity of estimates. The one-way ANOVA, $F(3, 18) = 0.548$, $MSE = 0.02$, $p = .656$, $\eta^2 = 0.08$, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that kindergarten groups exhibited similar changes in estimation patterns. Furthermore, contrasting performance of the linear board game intervention group and the control group provided similar results, $t(18) = 0.342$, $p = .74$, *d* = 0.03. Contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(18) = 1.165$, $p = .259$, $d = 0.09$, although significance of difference increased.

1st grade experimental condition and change in mean percent absolute errors. The one-way ANOVA, $F(3, 27) = 2.465$, $MSE = 0.002$, $p = .084$, $\eta^2 = 0.22$, indicated slightly significant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that first grade groups exhibited similar changes in estimation skills. Furthermore, contrasting performance of the linear board game intervention group and the control group provided similar results, $t(11) = -1.825$, $p = .095$, $d = -0.04$ as did contrasting performance of the broken number line puzzle intervention group and the control group, $t(10) = 1.983$, $p = .076$, $d = 0.05$.

1st grade experimental condition and change in linearity of estimates. The one-way ANOVA, $F(3, 27) = 1.620$, $MSE = 0.02$, $p = .208$, $\eta^2 = 0.15$, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that first grade groups exhibited similar changes in estimation patterns. Furthermore, contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(10) = 0.632$, $p = .542$, $d = 0.04$. Contrasting performance of the linear board game intervention group and the control group provided somewhat different results, $t(7) = -2.134$, $p = .069$, $d = -0.13$; differences in change in linearity appear more significant between the first grade linear board game intervention and control groups.

2nd grade experimental condition and change in mean percent absolute errors. The one-way ANOVA, $F(3, 38) = .659$, $MSE = 0.005$, $p = .582$, $\eta^2 = 0.05$, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that second grade groups exhibited similar changes in estimation skills. Furthermore, contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(38) = -0.173$, $p = .863$, $d = -0.01$. Contrasting performance of the linear board game intervention group and the control group provided similar results, $t(38) = -1.396$, $p = .171$, $d = -0.04$, although an increase in significance of differences in the contrast is indicated.

2nd grade experimental condition and change in linearity of estimates. The one-way ANOVA, $F(3, 38) = .732$, $MSE = 0.05$, $p = .539$, $\eta^2 = 0.05$, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that second grade groups exhibited similar changes in estimation patterns. Furthermore, contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(38) = -0.129$, $p = .898$, $d = -0.01$. Contrasting performance of the linear board game intervention group and the control group provided similar results, $t(38) = -1.371$, $p = .178$, $d = -0.13$, although an increase in significance of differences in the contrast is indicated.

Linear Board Game Discussion

Direction of differences indicated by the group contrasts were such that the linear board game group had slightly improved number line estimation results when compared to the control group, which was matched by ability based on relative performance on grade normalized achievement tests. Samples sizes, means, and standard deviations for the baseline and post assessment number line estimation tasks for these groups are summarized in Table 2; further discussion follows.

Table 2
Descriptive Statistics for Number Line Estimation Task Response Variables by Grade

	Linear Board Game Intervention *n* = 23		**Control** *n* = 23	
Time	Baseline	Post	Baseline	Post
MPAE	20.58 (19.78)	16.99 (17.88)	17.64 (13.77)	18.04 (14.31)
Linearity	0.60 (0.39)	0.68 (0.40)	0.62 (0.37)	0.60 (0.36)
Kindergarten	*n* = 5		*n* = 5	
MPAE (*SD)*	51.45 (11.28)	46.81 (13.67)	33.74 (7.11)	32.42 (11.38)
Lin (*SD)*	0.07 (0.08)	0.03 (0.04)	0.25 (0.14)	0.24 (0.05)
First Grade	*n* = 7		*n* = 8	
MPAE (*SD)*	14.25 (9.78)	11.79 (8.89)	19.75 (13.33)	21.07 (15.90)
Lin (*SD)*	0.66 (0.31)	0.77 (0.32)	0.54 (0.42)	0.52 (0.42)
Second Grade	*n* = 11		*n* = 10	
MPAE (*SD)*	10.58 (11.74)	6.75 (1.63)	7.9 (7.06)	8.4 (4.54)
Lin (*SD)*	0.81 (0.29)	0.93 (0.03)	0.87 (0.19)	0.86 (0.15)

Note. MPAE = Mean percent absolute error, Lin = variability explained by linear model

Overall group differences between experimental and control groups occurred in both mean percent absolute error and linearity. When controlling for grade differences, kindergarteners' estimation changes between experimental and control groups existed, but were not definitive; both groups had reduced absolute error, but also reduced degrees of linearity. Grade 1 appeared to have a greater difference in linearity of estimates but not accuracy, although the direction of mean absolute error improved in the intervention group compared to the control. Grade 2 groups were slightly different for both measures;

response means indicated improvement in the number line estimation task for students participating in the linear board game intervention when compared to those not in the number line activity group.

Broken Number Line Puzzle Discussion

Direction of differences indicated by the group contrasts provided mixed results for changes in number line estimation when the number line puzzle group was compared to the control group, which was matched by ability based on relative performance on grade normalized achievement tests. Samples sizes, means, and standard deviations for the baseline and post assessment number line estimation tasks for these groups are summarized in Table 3; further discussion follows.

Table 3

Descriptive Statistics for Number Line Estimation Task Response Variables by Grade

	Number Line Puzzle Intervention		**Control**	
	n = 25		*n* = 24	
Time	Baseline	Post	Baseline	Post
MPAE	18.14 (15.00)	17.73 (14.76)	16.89 (11.95)	16.36 (13.13)
Linearity	0.58 (0.39)	0.61 (0.38)	0.59 (0.36)	0.65 (0.36)
Kindergarten	*n* = 6		*n* = 6	
MPAE (*SD*)	34.18 (19.99)	32.12 (17.50)	29.99 (12.25)	33.20 (14.24)
Lin (*SD*)	0.30 (0.40)	0.20 (0.30)	0.29 (0.30)	0.29 (0.38)
First Grade	*n* = 8		*n* = 8	
MPAE (*SD*)	17.80 (10.15)	20.80 (13.31)	16.63 (8.87)	14.46 (7.35)
Lin (*SD*)	0.49 (0.38)	0.52 (0.37)	0.53 (0.35)	0.61 (0.31)
Second Grade	*n* = 11		*n* = 10	
MPAE (*SD*)	9.64 (5.83)	7.65 (1.54)	9.23 (6.36)	7.78 (3.48)
Lin (*SD*)	0.80 (0.27)	0.90 (0.06)	0.81 (0.27)	0.90 (0.09)

Note. MPAE = Mean percent absolute error, Lin = variability explained by linear model

Overall group differences between experimental and control groups in both mean percent absolute error and linearity were insignificant. When controlling for grade differences, kindergarteners' estimation changes between experimental and control groups exist, but were not definitive; the experimental group had reduced absolute error while the control had an increase, but there was also evidence of reduced degrees of linearity. Grade 1 appeared to have a greater difference in linearity of estimates. Both in accuracy and linearity, the direction of differences indicated improvement in the estimations of the control group compared to that of the intervention group. Grade 2 groups were similar for both measures; response means indicated improvement in the number line estimation task for students participating in the number line puzzle intervention, but these changes were nearly equal to those not in the number line activity group.

Changes in Number Line Estimates Moderated by Mathematical Achievement

Participants scoring low on fall achievement measure and change in mean percent absolute errors. The one-way ANOVA, $F(3, 13) = 2.478$, $MSE = 0.002$, $p = .107$, $\eta^2 = 0.37$, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that each group exhibited similar changes in estimation skills. Contrasting performance of the linear board game intervention group and the control group indicated virtually no distinction between groups, $t(13) = -0.55$, $p = .957$, $d = -0.002$; there appeared to be little significance in differences when limiting the comparison to these groups. Contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(13) = 2.562$, $p = .024$, $d = 0.08$, however there appeared to be slightly more significance in differences when limiting the comparison to these groups.

Participants scoring low on fall achievement measure and change in linearity of estimates. The one-way ANOVA, $F(3, 13) = 2.268$, $MSE = 0.03$, $p = .129$, $\eta^2 = 0.34$, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that each group exhibited similar changes in estimation linearity. Furthermore, contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(4) = 1.731$, $p = .163$, $d = 0.28$. Contrasting performance of the linear board game intervention group and the control group indicated virtually no distinction between groups, $t(6) = 0.111$, $p = .915$, $d = 0.005$; there appeared to be little significance in differences when limiting the comparison to these groups.

Participants scoring average on fall achievement measure and change in mean percent absolute errors. The one-way ANOVA, $F(3, 43) = 2.162$, $MSE = 0.007$, $p = .106$, $\eta^2 = 0.13$, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that each group exhibited similar changes in estimation skills. Furthermore, contrasting performance of the linear board game intervention group and the control group provided similar results, $t(14) = -1.833$, $p = .088$, $d = -0.08$ as did contrasting performance of the broken number line puzzle intervention group and the control group which provided similar but somewhat less significant results, $t(24) = -1.029$, $p = .314$, $d = -0.03$.

Participants scoring average on fall achievement measure and change in linearity of estimates. The one-way ANOVA, $F(3, 43) = 1.651$, $MSE = 0.04$, $p = .192$, $\eta^2 = 0.10$, indicated insignificant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that each group exhibited similar changes in estimation linearity. Furthermore, contrasting performance of the linear board game intervention group and the control group provided similar results, $t(12) = -1.753$, $p = .104$, $d = -0.17$as did contrasting performance of the broken number line puzzle intervention group and the control group, $t(20) = -1.370$, $p = .186$, $d = -0.09$.

Participants scoring high on fall achievement measure and change in mean percent absolute errors. The one-way ANOVA, $F(3, 27) = 0.430$, $MSE = 0.002$, $p = .733$, $\eta^2 = 0.04$, indicated insignificant differences in change in mean percent absolute errors between the four experimental condition groups. This evidence demonstrates that each group exhibited similar changes in estimation skills. Furthermore, contrasting performance of the broken number line puzzle intervention group and the control group provided similar results, $t(7) = 0.393$, $p = .706$, $d = 0.01$. Contrasting performance of the linear board game intervention group and the control group provided similar results, $t(9) = -0.625$, $p = .548$, $d = -0.01$, however there appeared to be slightly more significance attributed to differences when limiting the comparison to these groups.

Participants scoring high on fall achievement measure and change in linearity of estimates. The one-way ANOVA, $F(3, 27) = 2.836$, $MSE = 0.008$, $p = .057$, $\eta^2 = 0.24$, indicated slightly significant differences in change in linearity between the four experimental condition groups. This evidence demonstrates that groups exhibited somewhat different changes in estimation linearity. Contrasting performance of the broken number line puzzle intervention group and the control group provided results indicating significant differences when limiting the comparison to these groups, $t(27) = 2.663$, $p = .013$, $d = 0.13$. Contrasting performance of the linear board game intervention group and the control group provided insignificant results, $t(27) = -1.135$, $p = .266$, $d = -0.05$.

Discussion.

The lack of significance reflected in the comparisons masks a few interesting indicators of trends, including evidence that each number line activity resulted in improved number line estimation responses most notably for those in the average achievement group, without consideration of grade. Differentiating time with materials for the low group and extending proportional relationships for the high group may produce positive effects for all ability groups and is worthy of further consideration.

Additionally, the significance of the differences in the high achievement group was indicative of improvement in linearity in the control group ($n = 8$) paired with a reduction in linearity for the number line puzzle intervention group ($n = 7$). Because these groups were small, one possible explanation for the unusual change in performance of students in the control relative to the experimental is the presence of a proportionally large number of kindergarten members (three in each group). The variability of kindergarten responses would dramatically affect the results of this group. Contrast the membership of these groups with the low achievement score groups, which were also small. In that case both the puzzle and control group had four members, there was one kindergartener in each of those groups.

CHAPTER 6

Discussion: Number Line Activities as Interventions

Three questions were addressed in this study (1) How accurately do students in kindergarten, 1st grade and 2nd grade estimate numbers for locations on a line given only endpoint values? (2) Which model (linear or logarithmic) best fits students' mental representations of the number line? The third – and primary research question concerns intervention: (3) Does utilizing physical models shift students' mental representations of number lines to be more linear and less logarithmic?

Related to these questions there were five hypotheses for this dissertation: a) absolute error rates and linearity vary by grade level, b) absolute error rates will remain consistent over a four week period in the absence of direct intervention, whereas absolute error rates will decrease following 3 ten minute number line activity interventions, c) linearity of mental number lines will remain consistent over a four week period in the absence of direct intervention, d) linearity will increase following 3 ten minute number line activity interventions; furthermore, dependence on a logarithmic model will decrease, and e) dynamic number line activities will provide improved understanding compared to more passive representations; puzzles will yield stronger results than games and missing number opportunities will yield stronger results than activities in which all values are given. Hypothesis a) directly relates to questions (1) and (2) and addresses age differences in mental number line representations. The primary research question addressing change in mental representations due to intervention is addressed in hypotheses b) through e).

Absolute error rates and linearity vary by grade.

Absolute error rates and linearity were shown to vary by grade level; as school experience (grade) increased absolute error decreased and linearity increased. The baseline number line estimation task provided substantive evidence which supports the existing research by Ashcraft and Moore (2011), Booth and Siegler (2006, 2008), and Siegler and Booth (2004) and the log-to-lin shift model. However, evidence provided by naïve participants may be interpreted in the light of the segmented linear model (Moeller, 2009) and accumulator model (Ebersbach et al, 2008).

All three models appear equally insufficient to address the variability with which emergent learners mark estimates in the number line estimation task. There appears to be a progression from arbitrary end anchors, to magnitude related end anchors, with a subtle shift to near end anchors, before estimations appear to become less random. A few individuals seem to randomly mark estimates with no consideration of the given end anchors, which may indicate the most naïve understanding of number lines and possibly magnitude in general.

Additionally, the related hypothesis that mental number lines remain consistent in the absence of direct intervention was also supported. In all experimental control groups, with the exception of the high ability puzzle group described in Chapter 7, the baseline and post number line estimation tasks were not significantly different.

Changes in understanding from baseline to post assessment.

Number line activities generally resulted in reduced mean percent absolute errors and increased linearity, however the improvements were rarely statistically significant.

Contrast this with the linear board game research of Ramani and Siegler (2011), Siegler and Ramani (2009), and Whyte and Bull (2008), who provided evidence that this activity produces significant results. The lack of statistical significance paired with the direction of the findings indicates a need for further study.

The changes to the linear board game implemented in the current study (increased magnitude of numbers and student use of missing number strategies) were differences which increased cognitive load; this required more time and practice to internalize. Ashcraft and Moore (2012) saw similarly fragile understanding when measures changed from 0 to 100 number line estimation to 0 to 1000 estimations among school children; whereas college students maintained linearity for both trials. The apparent regression of learning is tentatively explored with activity-effect relationship (AER), the conceptual framework used by Tzur and Lambert (2011). Developed through consideration of Vygotsky's Zone of Proximal Development and Siegler's Overlapping Waves Model, the AER stage of learning suggests that mathematical conceptions are invariant anticipation a learner calls up to accomplish goals in familiar situations. The limited exposures to the parallel tasks and assessments in this study provided unstable conceptions; time and repeated experiences will solidify the relationship between the game or puzzle activity and the number line estimation task. AER, according to Tzur and Lambert, can reflect the learning that often follows completion of prompts when subjects notice that they did not need to complete an activity to find a solution. In the 2011 study an example is provided of a child counting a quantity of items that are then rearranged; the child recounts and soon after realizes they did not need to recount to provide the correct response. Repeated use and identity discovery makes the developed AER accessible to the child. Thusly, the hypothesis regarding greater effect of dynamic puzzles and games, as well as the missing value treatment on learning is yet to be determined. The complexity of representational change reflecting learning indicates extended time is essential to reach meaningful conclusions.

Another aspect this uncovers is the difference in presentation between the number line activities and the number line estimation task. The activities, although having some missing values, included some consecutive numbers and some hash marks which indicated the presence of unlabeled numbers, whereas the estimation task only had the endpoints, 0 and 100, and an empty line. Providing students with one number line on which to mark all values, rather than a new one for each number prompt may encourage magnitude discovery and demonstrate existing, but unstable, understanding particularly that of students with emergent numeracy skills.

Limitations of this study including possible confounds.

Study participants included a convenience selection of experimental and control groups by classroom, however the effect did not show itself in the board game intervention/control comparison. There were differences between the game and puzzle interventions; this provides evidence that intervention groups in the same grade which were taught in the same classroom did not contribute a significant confound.

In this field study sample size was limited by the population at the educational site. All students in kindergarten, 1st grade, and 2nd grade participated in the study, but the number of students in each cell was significantly smaller than the ideal. To improve reliability future studies should include samples from a variety of sites or settings to increase the population in each cell.

Variability in field collection of data and participation in the intervention activity.

Field research provides rich data because it is authentic, but with this richness there are limitations for generalization. Participants miss school due to illness, vacations, and appointments, school is cancelled or schedules are changed due to extreme weather, and teachers have varying scheduling needs. Flu season caused the loss of some students' data because they missed the baseline assessment. Included in the results but not explained are the following attendance issues: a) two kindergarten students, one in each activity, missed the third intervention, b) two first graders missed the first and second linear board game intervention, but were provided extra time during the third, and c) one second grader missed the first puzzle intervention. Additionally, first graders were simultaneously being pulled for interventions by a tutor, although they completed each prescribed intervention, it was with different group-mates in at least one of the three sessions. Of special note, those students experiencing the tutoring were students in need of academic support; in an effort to provide learning experiences, the dual obligation somewhat undermined the routines and procedures established for the study. The first grade subgroups had added variability, particularly for the students with achievement scores below average, compared to kindergarten and grade 2.

Time was held constant during the intervention sessions, which caused some groups to proceed further in session one. One linear board game kindergarten group completed the first version of the task (1-10 game with no missing values) where the other groups (one linear board game and two puzzle groups) also completed the second version with missing values; the remaining interventions were completed at the same rate by all kindergarten groups. The first grade groups all completed the same number of versions in each session. Grade two had the greatest variability in session 1; one group completed version one of the linear board game, while the other two linear board game groups also completed version two. The three number line puzzle groups completed three versions of the puzzle in the same amount of time. This variability in grade two was only present during intervention session one; the remaining sessions all groups completed three versions of their respective activity.

Time needed to clarify activity directions during intervention session one was different between groups and was the primary cause of the version completion variability. Because time with students was limited, it was held constant as an experimental control and as a practicality. The experimental significance of these departures from the process is important; however recognition of the reality of classroom time commitments for and from students is generally variable between classrooms and limited and must be considered in the context of the findings of this study.

Time was also a factor which may have limited the strength of the results. Siegler and Ramani (2009) detected children’s generation of linear representations after one hour of playing the linear board game. In this study children played for a total of 30 minutes and elicited results in the same direction, but less significant. More student time with the intervention materials is required to effectively assess the relative strengths of the activities.

CHAPTER 7

Conclusions

Early numeracy skills of young children are related to mathematics achievement in future grades (e.g., Geary, 2011 and Jordan et al, 2009). Early education generally focuses on one-to-one correspondence, counting, and number writing (Wright, Martland, Stafford, & Stanger, 2006) none of which sufficiently addresses linear magnitude. Linear games are a valuable means for exploring and practicing number relationships in an effort to learn or improve mental representations (Siegler & Ramani, 2009). Game play, which includes puzzles, provides multiple cues to form a robust linear representation. These cues include spatial, temporal, kinesthetic, verbal, and auditory magnitude relationships, if oral rehearsal and manipulation are part of the game (Whyte & Bull, 2008).

The present study included activities with missing information to produce hypotheses testing structures. Patterns are a strong tool for understanding numbers and mathematics. NCTM states that students “through experiences in school, should become more skilled in noticing patterns in arrangements…and in using patterns to predict what comes next in an arrangement” (2000, p. 91). Requiring prediction of missing numbers is an active teaching strategy used to increase engagement and attention, while simultaneously increasing the quality of teacher - student and student – material interaction through feedback, in this case regarding correct or incorrect missing number hypotheses.

Additionally, the number line puzzles provided opportunities for students to arrange the number line segments to produce a correct pattern. Sense-making is functionally important and the additional cues of end point arrows on the puzzles provided students with immediate feedback followed by the procedure of saying the sequence silently to check the entire puzzle. Students in the interventions used both puzzle format and counting strategies to ensure accuracy in puzzle completion. Existing errors were detected and corrected by students using these methods. On less than five occasions individual students were shown the completed puzzle to prompt a correction; this led to correction, at least once, without the student actually looking at the completed puzzle. This indicates a possibility that attention to feedback was not complete approximately two percent of the time, given that puzzles were completed over 250 times throughout the study.

Further research must be done to better understand the effects of the puzzle intervention on understanding magnitude and refining mental number line representation. Indicators are promising; both linear interventions resulted in change which was generally positive and had greatest impact for average learners. Ramani and Siegler (2011) found at the preschool level playing linear board games reduced the achievement gap, possibly remediating for lack of game and turn taking experience by lower achieving students. The students in this study had at least six months (kindergarten), often much more, of common educational experiences; did this produce the *average effect* indicated by the data? Adaptations to address the needs of students with both above and below achievement performance should be considered in future studies. This will inform educators' development of activities focused on processing and understanding the linear relationship of numbers for students at different levels of expertise in numeracy.

Upcoming research relative to linear activities must include separation of the missing value and continuous pattern number line activities, holding the activity constant. This will allow us to understand how the processing of missing values differs from the rehearsal of the complete pattern, providing clues for differentiation.

Additionally, puzzles and games must be further studied to determine the benefits of each. The added complexity of assembling complete number line segments compared to the practice provided by seeing the complete pattern will provide insights to conceptual development of numeracy, furthering our ability to differentiate activities for students with different levels of expertise.

A third area for consideration is the range of numbers provided in activities. The use of a 0 to 100 number line in the estimation task provides some guidance, but is the most effective method limiting experiences for preschool 0 to 10, primary grades 0 to 100, then intermediate grades 0 to 1000 until each level is mastered to ensure AER is solid and prevent regression? This also begins the conversation for introducing rational numbers, integers, and irrational numbers. The number line can clarify the relative magnitude of all numbers, modifying the endpoints and scale for instruction.

Finally, the reliance on endpoint anchors in the estimation task indicates a need for pairing estimation and number line activities with instruction regarding concepts of equality, greater than, and less than, particularly for naïve students. Some variations have been studied; Cohen and Blanc-Goldhammer (2011) varied the length of the number line and Karolis, Iuculano, and Butterworth (2011) added end point variability to ensure that decades were not a confound. Investigation of promising practices for these students must continue.

In conclusion, the prevalence of number in math requires developing a strong foundation in numeracy to which this research has contributed. Mathematics is an economic gate keeper which provides opportunities for those who have mastered it (i.e., Geary, 2011 and Ramani & Siegler, 2011). Mental number line representations are important referents for all students of mathematics. Our ability to improve students' early internal whole number representations will facilitate understanding of rational and irrational numbers. Providing appropriate opportunities to ensure that all learners understand and can use mathematical concepts is the goal of mathematics education and must be the focus of continued research.

References

Ashcraft, M.H., & Moore, A.M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*, 246-267.

Baroody, A. J., Bajwa, N.P., & Eiland, M. (2009). Why can't Johnny remember the basic facts? *Developmental Disabilities Research Reviews, 15*, 69-79.

Bay, J.M. (2001). Developing number sense on the number line. *Mathematics Teaching in the Middle School, 6*(8), 448-451.

Booth, J.L., & Siegler, R.S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*(6), 189-201.

Booth, J.L., & Siegler, R.S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016-1031.

Coffey, M.E. (2001). Irrational numbers on the number line: Perfectly placed. *Mathematics Teacher, 94*(6), 453-455.

Cohen, D.J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychon Bull Review 18*, 331-338, accessed online February 17, 2012.

Dixon, J.K. (2008). Tracking time: Representing elapsed time on an open timeline. *Teaching Children Mathematics, 15*(1) 18-24.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9- year old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99*, 1-17.

Geary, D.C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539-1552.

Jordan, N.C., Kaplan, D., Ramineni, C., & Locuniak, M.N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867.

Jordan, N.C., & Levine, S.C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15*, 60-68.

Karolis, V., Iuculano, T., & Butterworth, B. (2011). Mapping numerical magnitudes along the right lines: Differentiating between scale and bias. *Journal of Experimental Psychology: General, 4*, 693-706.

Kastberg, S. E., & Walker, V. (2008). Insights into our understandings of large numbers. *Teaching Children Mathematics, 14,*9 530-536.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H. (2009). Children's early mental number line: Logarithmic or decomposed linear? *Journal of Experimental Child Psychology, 103*, 503-515.

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

National Mathematics Advisory Panel. (2008). Foundations for success: the final report of the National Mathematics Advisory Panel. Washington, D.C.: U.S. Department of Education.

Northwest Evaluation Association. (2011). *Technical Manual for Measures of Academic Progress® (MAP®) and Measures of Academic Progress for Primary Grades (MPG)*. Portland, OR: Northwest Evaluation Association. Available online at http://www.nwea.org.

Pearson Education. (2012). *Aimsweb Technical Manual*. Bloomington, MN: Pearson Education, Inc. Available online at http://aimsweb.com.

Ramani, G.B., & Siegler, R.S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology, 32*, 146-159.

Siegler, R.S., & Booth, J.L. (2004). Development of numerical estimation in young children. *Child Development, 75,* 428-444.

Siegler, R.S., & Opfer, J.E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237-243.

Siegler, R. S., & Ramani, G.B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545-560.

Tzur, R., & Lambert, M. A. (2011). Intermediate participatory stages as zone of proximal development correlate in constructing counting-on: A plausible conceptual source for children's transitory "regress" to counting-all. *Journal for Research in Mathematics Education, 42*(5), 418-450.

Whyte, J.C., & Bull, R. (2008). Number games, magnitude representations, and basic number skills in preschoolers. *Developmental Psychology, 44*(2), 588-596.

Wright, R.J., Martland, J., Stafford, A.K., & Stanger, G. (2006). *Teaching number: Advancing children's skills & strategies* (2nd ed.). Thousand Oaks, CA: Sage Publications.