The Pennsylvania State University

The Graduate School

College of Education

PROMOTING EARLY NUMERACY SKILL GROWTH IN HEAD START CHILDREN

A Dissertation in

School Psychology

by

Erin E. Reid

© 2010 Erin E. Reid

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2010

The dissertation of Erin E. Reid was reviewed and approved* by the following:

James C. DiPerna
Associate Professor of Education
Professor-in-Charge of Graduate Programs in School Psychology
Dissertation Adviser
Chair of Committee

Mark T. Greenberg
Professor of Human Development and Family Studies and Psychology

Robert L. Hale
Professor of Education

Pui-wa Lei
Associate Professor of Educational Psychology

*Signatures are on file in the Graduate School

ABSTRACT

Although evidence suggests that Head Start has demonstrated some success in promoting certain child outcomes, such as vocabulary development and social skills, recent reports indicate that Head Start children are performing approximately a full standard deviation below national norms on standardized measures of early mathematics (U.S. DHHS, 2005; 2006a; 2006b; 2010). The purpose of this study was to evaluate the effectiveness of an early numeracy intervention designed specifically for use in Head Start classrooms. Ninety-six children from six Head Start classrooms (three intervention and three comparison) participated in the study. During a 13-week period, intervention teachers facilitated one whole- or small-group activity and one transition activity per day. In addition, intervention teachers provided mathematics activities for children to explore independently at a math center. Child outcomes were measured before and after the intervention using the Test of Early Mathematics Ability – 3rd Edition and the EARLI Numeracy probes. Teacher reports and classroom observations were used to assess implementation fidelity, quality of mathematics instruction in the classrooms, and teacher acceptability of the intervention.

Hierarchical regression analyses were run to determine if intervention condition was a significant predictor of the criterion variables when controlling for pretest scores and demographic characteristics. Overall, children who participated in the intervention made significant gains on the post-test measures. However, the gains were not significantly greater than those made by children in the comparison condition. Descriptive data indicated that intervention teachers spent more time facilitating mathematics activities and provided higher quality instruction than comparison teachers. Classroom observation and teacher report indicated the intervention teachers conducted the intervention as designed; however, the accuracy of the teacher report data was questionable due to several inconsistencies in reporting. Intervention teachers reported satisfaction with the intervention, but indicated ways in which the intervention could be improved.

TABLE OF CONTENTS

List of Tables…………………………………………………………………………….. vii

Acknowledgements xi

Chapter 1: Introduction…………………………………………………………………... 1

Chapter 2: Literature Review…………………………………………………………… 6

Views on Early Mathematical Competence…………………………………….......... 6

Number Sense……………………..…………………………………………………….. 7

Short-term Early Mathematics Interventions………………………………………… 10

Long-term Early Mathematics Interventions…………………………………………. 15

Conclusions and Rationale for the Current Study……………………………………. 19

Research Questions………………………………………………………………………. 21

Chapter 3: Method..………………………………………………………………………. 22

Participants……………………………………………………………………………… 22

Number*Fun* Early Intervention Program…………………………………………….. 26

Measures………………………………………………………………………………….. 28

Procedures……………………………………………………………………………….. 34

Design and Analysis…………………………………………………………………….. 37

Treatment of Missing Data in Outcome Measures……………………………......... 40

Chapter 4: Results...……………………………………………………………………....... 48

Number*Fun* Implementation Fidelity…………………………………………………....... 48

Characteristics of Early Mathematics Instruction…………………………………… 49

Teacher Acceptability of Number*Fun* Intervention………………………………….. 53

Child Outcomes…………………………………………………………………………. 54

Chapter 5: Discussion……………………………………………………………….……. 67

Summary of Key Findings Relative to Hypotheses.……………………………......... 67

Interpretation of Results Relative to Previous Research…………………………....... 71

Working Hypotheses for Outcomes…………………………………………………….. 72

Limitations……………………………………………………………………………….. 76

Future Directions for Research………………………………………………………….. 79

Conclusion……………………………………………………………………………….. 80

References………………………………………………………………………………….. 82
Appendix A: Head Start Child Outcomes Framework—Mathematics………………….. 96
Appendix B: National Council of Teachers of Mathematics Curriculum Focal Points and Connections/Expectations…………………………………………………….…...... 97
Appendix C: Pennsylvania Department of Education Logical-Mathematics Early Learning Standards………………………………………………………………………. 98
Appendix D: Teacher Recruitment Letter……………………………………………….. 100
Appendix E: Teacher Consent Form…………………………………………………….. 102
Appendix F: Parent Recruitment Letter: English………………………………………... 104
Appendix G: Parent Recruitment Letter: Spanish……………………………………….. 105
Appendix H: Parent Consent Form: English…………………………………………….. 106
Appendix I: Parent Consent Form: Spanish……………………………………………... 108
Appendix J: Skills Targeted by Number*Fun* Intervention………………………………. 110
Appendix K: Sample Number*Fun* Activity: Number and Operations…………………... 112
Appendix L: Sample Number*Fun* Activity: Geometry and Spatial Sense………………. 113
Appendix M: Sample Number*Fun* Activity: Patterns and Measurement……………….. 114
Appendix N: Internal Consistency and Item Analysis for EARLI Numeracy Probes…... 115
Appendix O: Classroom Observation of Early Mathematics Instruction……………….. 116
Appendix P: Early Mathematics Daily Activity Log……………………………………. 118
Appendix Q: Number*Fun* Daily Activity Log…………………………………………... 119
Appendix R: Number*Fun* Intervention Program Teacher Feedback……..………….…... 121
Appendix S: Number*Fun* Assessment Checklist………………………………………... 123
Appendix T: Descriptive Statistics for COEMI Items Averaged Across Activities……. 124
Appendix U: Teacher Ratings of Satisfaction with NumberFun Curriculum…………… 126
Appendix V: Descriptive Statistics for the Test of Early Mathematics Ability-3rd Edition Based on Original and Imputed Datasets…………………………………... 129
Appendix W: Descriptive Statistics for the Counting Aloud EARLI Numeracy Probe Based on Original and Imputed Datasets…………………………………………… 131
Appendix X: Descriptive Statistics for the Counting Objects EARLI Numeracy Probe Based on Original and Imputed Datasets…………………………………………… 133

Appendix Y: Descriptive Statistics for the Grouping EARLI Numeracy Probe Based on Original and Imputed Datasets…………………………………………………… 135
Appendix Z: Descriptive Statistics for the Measurement EARLI Numeracy Probe Based on Original and Imputed Datasets………………………………………… 137
Appendix AA: Descriptive Statistics for the Number Naming EARLI Numeracy Probe Based on Original and Imputed Datasets………………………………………… 139
Appendix AB: Descriptive Statistics for the Pattern Recognition EARLI Numeracy Probe Based on Original and Imputed Datasets…………………………………. 141
Appendix AC: Pooled Results of Reversal of Multiple Regression Analyses for TEMA-3……………………………………………………………………………….. 143
Appendix AD: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Counting Aloud……………………………………………………………... 144
Appendix AE: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Counting Objects…………………………………………………………. 145
Appendix AF: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Grouping………………………………………………………………………. 146
Appendix AG: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Measurement…………………………………………………………... 147
Appendix AH: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Number Naming……………………………………………………………….. 148
Appendix AI: Pooled Results of Reversal of Multiple Regression Analyses for EARLI Pattern Recognition…………………………………………………………. 149
Appendix AJ: Pooled Results of Multiple Regression Analyses for EARLI Counting Aloud with Classroom Added as a Predictor…………………………………….. 150
Appendix AK: Pooled Results of Multiple Regression Analyses for EARLI Counting Objects with Classroom Added as a Predictor…………………………………… 151
Appendix AL: Pooled Results of Multiple Regression Analyses for EARLI Grouping with Classroom Added as a Predictor……………………………………………. 152
Appendix AM: Pooled Results of Multiple Regression Analyses for EARLI Measurement with Classroom Added as a Predictor……………………………... 153

Appendix AN: Pooled Results of Multiple Regression Analyses for EARLI Pattern Recognition with Classroom Added as a Predictor…………………………………… 154

LIST OF TABLES

Table 1. Recent Definitions of Number Sense within Early Childhood Literature………………………………………………………………………………. 8
Table 2. Teacher Demographic Information…………………………………………….. 23
Table 3. Demographic Characteristics of Participants by Classroom…………………… 24
Table 4. Comparison of Demographic Characteristics Between Participants in Intervention and Comparison Conditions……………………………………….. 26
Table 5. Distribution of Missing Values in Child Mathematics Measures……………… 43
Table 6. Composite Scores for COEMI Domains and Fidelity Scale Averaged Across Observations……………………………………………………………………… 52
Table 7. Daily Average Number and Length of Mathematics Activities by Classroom… 53
Table 8. Total Teacher Satisfaction Ratings of Number*Fun* Intervention by Domain….. 53
Table 9. Teacher Ratings of Number*Fun* Activity Acceptability………………………. 54
Table 10. Pooled Means and Standard Errors of Early Mathematics Outcomes by Intervention Condition…………………………………………………………… 55
Table 11. Pooled Means and Standard Errors of Early Mathematics Outcomes by Classroom………………………………………………………………………… 56
Table 12. Intraclass Correlation Coefficients (ICC) of Mathematics Measures Averaged Across Imputed Datasets……………………………………………………… 58
Table 13. Pooled Results of Multiple Regression Analyses for Test of Early Mathematics Ability – 3rd Edition……………………………………………….. 59
Table 14. Pooled Results of Multiple Regression Analyses for EARLI Counting Aloud…………………………………………………………………………….. 60
Table 15. Pooled Results of Multiple Regression Analyses for EARLI Counting Objects……………………………………………………………………………. 61
Table 16. Pooled Results of Multiple Regression Analyses for EARLI Grouping…...…. 62
Table 17. Pooled Results of Multiple Regression Analyses for EARLI Measurement…. 63
Table 18. Pooled Results of Multiple Regression Analyses for EARLI Number Naming…………………………………………………………………………… 63

Table 19. Pooled Results of Multiple Regression Analyses for EARLI Pattern Recognition………………………………………………………………………. 64
Table 20. Pooled Results of Multiple Regression Analyses for Test of Early Mathematics Ability – 3rd Edition with Classroom Added as a Predictor………….. 66
Table 21. Pooled Results of Multiple Regression Analyses for EARLI Number Naming with Classroom Added as a Predictor………………………………………………. 66

ACKNOWLEDGEMENTS

Words cannot adequately express my gratitude to my committee chair, and mentor, Dr. James DiPerna. Thank you for all the time and energy you have devoted to molding me into the scholar I am today. I am eternally grateful for all of the opportunities I have had under your wing. Without your enthusiasm, support, and encouragement my graduate school experience would not have been as rewarding.

I would like to thank my committee members, Dr. Pui-wa Lei, Dr. Robert Hale, and Dr. Mark Greenberg who offered their expertise and guidance as I completed this project. Thank you to all of the Lebanon County Head Start teachers, staff, parents, and children who participated in this study. Many thanks to the School Psychology doctoral students who helped me collect the data necessary to produce this paper.

I also would like to thank Pete for his support and patience. You picked up the slack when I was locked away in the office, nose deep in drafts and revisions. Thank you for enduring the long nights and occasional emotional outbursts. You have taught me so much through this process; for example, you are the better cook and dishwasher. Keep up the good work!

Finally, this dissertation is dedicated to Jacob, my source of inspiration, determination, and perseverance. Without you, I would not be where I am today and would not have accomplished all that I have these past 13 years. We have made many sacrifices as I pursued this dream, but please know that I did it all for you, so that you would be as proud of me as I am of you.

The project was supported by the U.S. Department of Health and Human Services Administration for Children and Families (Award No. 90YR000501).

Chapter 1

Promoting Early Numeracy Skill Growth in Head Start Children

The consequences of poorly developed mathematical skills are far reaching, impacting overall educational level and employment (Mazzocco & Thompson, 2005). Considering that high school graduates with greater mathematics skills are more likely to complete college (Adelman, 1999; Dougherty, 2003; Evan, Gray, & Olchefske, 2006) and earn more (Mitra, 2002; Murnane, Willett, Duhaldeborde, & Tyler, 2000), study after study reports that U. S. high school graduates are deficient in mathematics skills in comparison to other nations (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008; Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2007). One report on mathematics and science achievement stated, "Nationally, 22% of all college freshmen fail to meet the performance levels required for entry-level mathematics courses and must begin their college experience in remedial courses" (Evan et al., 2006, p. 15).

With low mathematics proficiency potentially related to poorer outcomes in adulthood, the importance of fostering mathematical abilities in school-age children becomes more apparent, particularly in light of findings from the National Assessment of Educational Process (NAEP). Results of NAEP in 2005 indicated that 77% of twelfth grade students performed below proficiency on the mathematics measure (Grigg, Donahue, & Dion, 2007). In 2009, 66% of eighth grade students and 61% of fourth grade students performed below proficiency in mathematics (National Center for Educational Statistics, 2009). These national evaluation efforts highlight that a sizable proportion of U. S. children do not demonstrate a firm understanding of mathematical concepts appropriate for their grade level. Also apparent are the disparities in mathematics achievement for children living in poverty, a group historically at-risk for experiencing academic difficulties. An indicator of income level in NAEP is eligibility for free or reduced-price lunches. For eighth graders, a gap in scores of 18 to 29 scale points (reduced-price and free lunch eligible, respectively) was observed. In fourth grade, the gap was slightly less, 15 to 24 scale points (reduced-price and free lunch eligible, respectively).

These income level disparities in mathematics skills are evident before children enter formal school (Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Jordan, Kaplan, Oláh, & Locuniak, 2006; National Science Board, 2010), even as early as preschool entry (Malofeeva, Ciancio, & Day, 2001; Saxe, Guberman, & Gearhart, 1987; Starkey, 1992). Using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Denton and West (2002) observed that kindergarten children living in poverty performed far under the national average on measures of mathematics achievement, and this persisted into first grade. More recent research suggests that the level of mathematics skills at kindergarten entry is relatively stable beyond first grade, even into eighth grade. (Duncan, Dowsett, Claessens, Magnuson, Huston, Klebanov, et al., 2007; Halberda, Mazzocco, & Feigenson, 2008; National Science Board, 2010). Young children who lack fundamental numeracy skills tend to have difficulty with more complex mathematical concepts when they begin formal schooling. This, in turn, causes them to remain behind their more knowledgeable peers (Geary, Hoard, & Hamson, 1999) and, in some instances, fall further behind their peers (Aunola et al., 2004; Williamson, Appelbaum, & Epanchin, 1991). If not addressed, these early differences may turn into a disability later on in their schooling (Case, Griffin & Kelly, 1999).

Because a gap in mathematics achievement becomes apparent early, preschool provides an opportune time to intervene with at-risk children. Fueled by the National Education Goals Panel's (1999) declaration that all children will enter school ready to learn, there has been increased interest in the promotion of early academic skills, particularly early literacy skills. Despite evidence that young children living in poverty enter school with fewer early numeracy skills than their middle-income peers (Aunola et al., 2004; Jordan et al., 2006; Klein & Starkey, 2004), these skills have received less attention than early literacy skills (Graham, Nash, & Paul, 1997).

Many preschool classrooms do not succeed in promoting early numeracy skills (Sarama, DiBiase, Clements, & Spitler, 2004). Both Arnold, Fisher, Doctoroff, & Dobbs (2002) and Peters (1998) demonstrated that numeracy skills do not need to be taught in a didactic manner in order to produce positive gains in young children's numeracy skills. In fact, many preschool children already engage in some form of mathematical activity (e.g., classification, comparison, enumeration, and creating patterns) during free play without adult direction (Seo & Ginsburg, 2004). However, the more preschool teachers talk with their students about early numeracy concepts, the more knowledge the children gain by the end of the school year (Kilbanoff, Levine, Huttenlocker, Vasilyeva, & Hedges, 2006). By systematically weaving numeracy activities into classroom routines (e.g., snack, transition, free play) and encouraging children to think like mathematicians through discussion, children can discover and develop fundamental numeracy skills needed for later success in mathematics during their elementary years (Arnold et al. 2002; Kilbanoff et al., 2006).

The nation's largest targeted intervention program, Head Start, has the goal of helping children living in poverty "catch up" with their more advantaged peers by the start of kindergarten (ACF, 2002). However, reports of its effectiveness in improving early math performance have been discouraging. Within the past decade, the U. S. Department of Health and Human Services (U.S. DHHS) has undertaken two ongoing longitudinal studies to quantify the impact of Head Start on the lives of the children it serves. The Family and Child Experiences Survey (FACES; U.S. DHHS, 2006a) examines child, family, and program characteristics and child outcomes resulting from enrollment in Head Start. The Head Start Impact Study (U.S DHHS, 2005; 2010) is a randomized longitudinal study investigating the outcomes of children participating in Head Start compared to children who were eligible for Head Start but were not enrolled.

Findings from FACES 2000 (U.S.DHHS, 2006a) indicated that 4-year-old children entering Head Start performed nearly a standard deviation below the national norm on an early mathematics measures. By the end of the same academic year, children had gained an average of 1.2 standard score points on the measure. Although the gain from fall to spring was statistically significant (most likely due to the large sample size), the effect size obtained (d = .08) indicated that the gain was negligible. Similar results were observed with FACES 2003 (U.S. DHHS, 2006b). Children enrolled in Head Start during the 2003-04 school year made an average gain of 1.5 standard score points on an early mathematics measure from the fall to the spring. By the end of kindergarten, Head Start graduates had gained approximately 6 standard score points on the measure, however, the children continued to perform below the national average. These results are consistent with the findings of the Head Start Impact Study. Four-year-old children enrolled in Head Start did not make significantly greater gains in early mathematics skills than children who were eligible for Head Start but did not receive services (U.S DHHS, 2005; 2010).

These reports indicate that children enrolled in Head Start are not making sufficient gains in early numeracy to catch up to their peers, which may lead them to struggle in acquiring elementary mathematics once they enter formal schooling. Thus, children may not be receiving a “head start” when it comes to early numeracy. Fortunately, individual differences in early numeracy are detectable at early ages. More importantly, research has shown that children, particularly those behind their peers, make significant and meaningful gains in numeracy skills after participating in early mathematics interventions (e.g., Pasnak, Greene, Fergusom, & Levit, 2006; Sophian, 2004; Young-Loveridge, 2004). Thus, identifying young children who are at-risk for future mathematics difficulties and intervening before these children enter school might alter the path of their expected growth trajectory into one that is more favorable.

The purpose of this study was to evaluate the efficacy of an early numeracy intervention program in Head Start classrooms. The program builds on previous research (e.g., Arnold et al., 2002; Pagani, Jalbert, & Girad, 2000; Peters, 1998; Starkey, Klein, & Wakeley, 2004; Young-Loveridge, 2004) and is standards based. Three Head Start teachers were asked to facilitate two small or whole group mathematics activities a day and to set up a math center for children to engage in specific activities independently while three other Head Start teachers continued to use the standard curriculum. Multiple early mathematics measures were completed by all children. Hierarchical regression analyses were conducted to determine if children in the intervention condition significantly outperformed children in the comparison condition. Additional information was collected on the quality of mathematics instruction, intervention fidelity, and teacher satisfaction with the intervention. This information was used to aid interpretation of the child outcomes.

Chapter 2

Literature Review

Views on Early Mathematical Competency

The goal of mathematics education at all grade levels is to support mathematical competence in students (Baroody, Lai, & Mix, 2006). Like so many other constructs in the fields of education and psychology, defining what constitutes mathematical competence is challenging due to the varying perspectives of experts in the field and the complexity of the underlying concepts. As Griffin (2003) writes,

> We can all recognize math competence when we see it and, given a bit of experience with individuals, we are all likely to know who has it and who doesn't…. When it comes to defining what math competence is…I suspect that many of us, including the math teachers and educators among us, might be at a loss (p. 3).

Within the last 20 years, more attention has been paid to children's mathematical competence during the preschool years. Prior to that, Piaget's (1965) views that preschool children were unable to engage in abstract and logical thinking and that cognitive development could not be accelerated heavily influenced the field of early childhood education. The result was little time devoted to mathematics related activities in early childhood settings. In fact, some argued that introducing advanced concepts before a child was cognitively ready could do more harm than good (Elkind, 1981).

While Piaget's constructivist views persist today in many early childhood classrooms, the notion that young children can develop early mathematical knowledge is widely accepted. There is evidence that young children demonstrate a wide range of numeracy skills, some of which are evident in infancy. Children as young as two or three can count small groups of objects, add or subtract with small numbers, and extend simple patterns, with little or any adult assistance (Ginsburg, Cannon, Eisenband, & Pappas, 2006). By the end of the preschool years, typically developing children (primarily middle-class) can recite number words 1-10 reliably and even continue to 100, count groups of objects, perform simple addition and subtraction problems, and identify geometric shapes (National Mathematics Advisory Panel, 2008a). Many of these skills are acquired without formal instruction.

Number Sense

The "informal" mathematical knowledge that preschoolers gain from everyday interactions with parents, caregivers, siblings, and/or peers has been characterized as number sense (Gersten & Chard, 1999). As with the term *mathematical competence*, there is a lack of consensus on a definition for number sense. Table 1 presents a sampling of recently published definitions offered by researchers in various fields.

Table 1. *Recent Definitions of Number Sense within Early Childhood Literature*

Source	Definition
Berch (2005, p. 334)	Permits one to achieve everything from understanding the meaning of numbers to developing strategies for solving complex math problems; from making a simple magnitude comparison to inventing procedures for conducting numerical operations; and from recognizing gross numerical errors to using quantitative methods for communicating, processing, and interpreting information.
Case (1998, p. 1)	Students with good number sense can move seamlessly between the real world of quantities and the mathematical world of numbers and numerical expressions….They can represent the same number in multiple ways depending on the context and purpose of this representation. They can recognize benchmark numbers and number patterns….They have a good sense of numerical magnitude and can recognize gross numerical errors…Finally, they can think or talk in a sensible way about the general properties of a numerical problem or expression – without doing any precise computation.
Gersten & Chard (1999, pp. 19-20)	A child's fluidity and flexibility with numbers, the sense of what numbers mean, and an ability to perform mental mathematics and to look at the world and make comparisons.
Jordan et al. (2006, p. 154)	The ability to subitize small quantities, to discern number patterns, to count, and to perform simple number transformations.
National Council of Teachers of Mathematics (2000, p. 32).	Fundamental understanding of, and proficiency with, counting, numbers, and arithmetic, s well as an understanding of number systems and their structures.
National Mathematics Advisory Panel (2008b, p. 27)	An ability to immediately identify the numerical value associated with small quantities…, a facility with basic counting skills, and a proficiency in approximating the magnitude of small numbers of objects and simple numerical operations.

In addition to the debate on the definition of number sense, there is debate as to the origin of these informal skills. Researchers in the cognitive sciences hold the view that number sense results from an innate ability to perceive numerical quantities (Dehaene, 1997). Evidence in support of this view is often found in studies conducted with infants. For instance, some research suggests that infants can discriminate between sets with different numerosities (e.g., Lipton & Spelke, 2003; Starkey, Spelke, & Gelman, 1990), that they can discriminate between equal and unequal quantities (e.g., Brannon, 2002; Cooper, 1984), and that they have a rudimentary ability to add and subtract small quantities (Van de Walle, Carey, & Prevor, 2000; Wynn, 1992). Educational researchers, on the other hand, hold the view that, while there may be an underlying biological mechanism, number sense comprises a complex understanding of numerical concepts that results from learning experiences (Berch, 2005).

Referring back to Table 1, despite the broad and abstract definitions of number sense related terminology, there is some consistency in the types of skills identified, such as understanding relationships between numbers and recognizing quantities. Jordan et al. (2006) identified five key components of number sense that may differentiate young children at-risk for developing mathematics difficulties. These components include: counting; number knowledge (e.g., recognizing that “4” is larger than “2”); number transformation (e.g., adding or subtracting by “1”); estimation (e.g., guessing the number of objects in a group); and number patterns (e.g., knowing that “8” is the missing number in the sequence: 2-4-6-?).

Individual differences in number sense or early mathematical skills are evident early. This may be due, in part, to differences in home environments prior to school entry. Parents typically emphasize literacy skills more than numeracy skills in their interactions with children (Blevins-Knabe & Musun-Miller, 1996; LeFevre, Clarke, & Stringer, 2002). In addition, parents attitude toward and experiences with mathematics influence the extent to which they engage their children in early mathematics activities (Skwarchuk, 2009). Socio-economic status (SES) related differences in the number of mathematics activities, as well as range of skills targeted, have been observed. Children from lower-SES homes tend to be exposed to fewer activities (Starkey, Klein, Chang, Dong, Pang, & Zhou, 1999). Given this information, it appears logical that preschool programs can play an important role in supporting young children's early numeracy development; however, early mathematics is not frequently taught in preschool classrooms (Balfanz, 1999; Saracho & Spodek, 2008).

To address the lack of quality mathematics instruction in preschool classrooms, several research teams have developed interventions to support the acquisition of early mathematics skills. These interventions ranged from short-term 6-week programs to interventions that span the academic year. Some interventions are conducted by research assistants working one-on-one with children, whereas other interventions are conducted in classrooms by teachers. These interventions and their impact on early mathematic outcomes are described below. Short-term interventions will be discussed first, followed by long-term interventions.

Short-term Early Mathematics Interventions

Math is Everywhere (Arnold et al., 2002). Arnold et al. conducted a 6-week intervention in Head Start classrooms serving 3- and 4-year-olds that targeted the following skills: counting, recognizing and writing numbers, one-to-one correspondence, comparison, change operations, and understanding numbers and quantity. Activities included books, music, games, discussion, and group projects. Four teachers implemented the activities in their classroom while four other classrooms did not. The researchers provided teachers with a 2-hour training and a book of 85 mathematics activities. For the first 3 weeks, teacher facilitated one activity during circle time. During the last 3 weeks, teachers conducted two activities during transition periods and/or mealtime.

Researchers collected data on children's mathematics skills before and after the intervention using the Test of Early Mathematics Ability-2nd Edition (TEMA-2: Ginsburg & Baroody, 1990). They also used surveys to assess children's interest in math and teachers' attitudes toward math. Children who participated in the activities gained an average of 3.67 (raw score) points, whereas children in the control classrooms gained .84 points. Comparison of gain scores across groups resulted in a significant intervention effect, $t(101) = 4.64$, $p < .01$, $d = 1.21$. The research team observed significant gender and ethnic differences in response to the intervention. That is, boys in the intervention condition gained nearly 2 more raw score points than intervention girls. In addition, African-American and Hispanic children gained approximately 3 and 4 more raw score points, respectively, than Caucasian children (Caucasians were less represented in the sample). Compared to control classrooms, significant gains in children's interest in math and teachers' attitudes toward math were found, as well.

Number Worlds (Pagani et al., 2000). Number Worlds is guided by five principles (Griffin, 2004). First, learning should build upon children's existing knowledge. Second, instruction should follow the natural progression of mathematical skill development. Third, instruction should support conceptual understanding in addition to computational fluency. Fourth, learning occurs through exploration and dialogue. Fifth, instruction introduces children to the various representations of number. The program targets children who are age eligible for kindergarten the following year. Teachers utilize games and activities to promote skills such as: identifying colors and shapes; verbal counting; taking away and adding to; comparing quantities; recognizing dimensions (e.g., big, medium, small); knowledge of the number line; estimation, and understanding fractions. Teachers participated in a 2-day training conducted by research staff. The training consisted of lectures, "interactions," and review of each unit of the program. For a period of 16 weeks, intervention teachers facilitated activities three times a week for 30 minutes per session. The authors did not provide information regarding the setting in which these activities took place (i.e., small- or whole-group). Teachers were instructed to work through the program at a pace that allowed the children to master the concepts and keep their interest.

All participating children completed the Number Knowledge Test (Okamoto & Case, 1996) before and after the intervention. Researchers reported that the intervention had a significant, yet small, impact on post-intervention number knowledge ($\eta_p^2 = .016$). Treatment condition accounted for nearly 2% of the unique variance in post-test scores. (Pretest scores and linguistic background were included as covariates). When the researchers compared high-dosage to medium-dosage to no-dosage implementation, they found that the intervention had a significant, yet small, effect on number knowledge preschool post-test scores ($\eta_p^2 = .019$).

Linear number board games (Ramani & Siegler, 2008). This short-term intervention was aimed at improving the number line estimation, numerical magnitude comparison, counting to 10, and number recognition of 4-year-old children. Over a 2-week period, research assistants worked one-on-one with children in Head Start classrooms (mean age 4-years, 9-months). During each of four sessions (15 – 20 minutes each) children played a specially designed board games with the assistant. The board game consisted of a spinner with numbers 1-10 printed on it and a board with 10 colored squares, lined up horizontally, each numbered consecutively from 1-10. On their turn, children spun the spinner and verbally identified the number. They then moved their token on the board that number of spaces while naming the numbers in the square. For instance, if the child was on the number 1 square and spun a 3, the child was to respond, "2, 3, 4." Correction was provided when necessary. Children in the control group played a similar game, only with colors instead of numbers. The board games were played 20 times over 4 sessions, each session lasting 15-20 minutes.

Children completed an assessment created by the researchers specifically for this study before, immediately after, and 9 weeks after the intervention. The assessment consisted of 4 subtests (number identification, numerical magnitude comparison, counting, and number line estimation). The researchers found large significant main effects for age, $F(5, 116) = 11.29$, $p < .001$, $\eta_p^2 = .33$; for testing session, $F(10, 111) = 11.41$, $p < .001$, $\eta_p^2 = .51$; for intervention condition, $F(10, 111) = 6.23$, $p < .001$, $\eta_p^2 = .21$; and a large, significant condition by session interaction effect, $F(10, 111) = 6.76$, $p < .001$, $\eta_p^2 = .38$.

Mathematics books and games (Young-Loveridge, 2004). Young-Loveridge investigated the utility of commercially available books and games as a means of increasing the numeracy skills of 4- and 5-year-old children in New Zealand. The objectives of the intervention were to develop knowledge of number word sequences, to develop proficiency with enumeration, to gain experience forming collections of objects, and to recognize number patterns and numerals. Participants of this study were selected based on their performance on a researcher-developed measure of numeracy skills; that is, children with a lower performance on the measure were identified as participants for the intervention. Teaching specialists pulled pairs of students out of class for 30 minutes a day, 5 days a week, for 7 weeks. Each session began with the specialist reciting a number rhyme and then reading a number story. During the story, children identified written numerals, checked quantities in the stories using counting, and made prediction about which quantities would come next. Next, children played a familiar board game for a brief amount of time. The specialist then introduced a new intervention game. New games focused on counting, one-to-one correspondence, number patterns, and number sequences using board games, dice, playing cards, and manipulatives.

Children completed the same numeracy assessment used for the screening pre-test, with some additional items. Follow-up assessments were conducted 6- and 15- months after the conclusion of the intervention. Again, the same measure was used, but additional items were added. At the conclusion of the intervention, children who participated in the intervention had significantly higher gains on the numeracy post-test than children who did not receive the intervention (d = 1.99). These students' gains persisted at 6 months (d = 1.12) and at 15 months (d = .50). At 15 months, however, the 17- point intervention advantage at the first post-test had decreased to 6 points. The researchers attributed this decrease to the fact that the curriculum used in kindergarten and the following grade did not support the skills gained by the students during the intervention and thus, allowing their comparison peers to gradually catch up.

Summary. Despite different approaches to promoting early mathematics skills and the short duration of the interventions, the aforementioned studies reported being relatively successful in improving early mathematics skills. Three of the four interventions led to large effects. Given that children in the intervention conditions participated in 30-150 minutes of mathematics activities per week, this is not surprising in light of literature suggesting that little mathematics instruction occurs in preschool classrooms (Balfanz, 1999; Saracho et al., 2008). The more time children were engaged in mathematics activities, however, did not appear related to the size of these effects. For instance, the intervention with the second highest time in minutes per week (i.e., Pagani et al., 2000), demonstrated a small effect.

Based on two factors, it is possible that the effects observed in these studies were overestimated. First, only one study (i.e., Arnold et al., 2002) reported reliability information for the measures used to assess child outcomes. Two of the studies that did not report reliability or validity information on the assessments (i.e., Ramani et al., 2008; Young-Loveridge, 2004) used measures developed specifically for evaluating the intervention. Second, two studies (i.e., Arnold et al., 2002; Young-Loveridge, 2004) conducted their analyses using mean gain for each condition scores. Using gain scores is problematic because the change in scores may be associated with changes in the situational factors surrounding the testing situation instead of changes in the skill being measured (Matton, Vautier, & Raufaste, 2009).

Long-term Early Mathematics Interventions

Building Blocks (Clements & Sarama, 2008). The skills targeted by the Building Blocks curriculum include: number (counting, number recognition, matching, subitizing, comparing number, sequencing, and arithmetic); geometry (shape identification, comparing shapes, representing shape, composing shape, and measurement); and patterning. Intervention teachers were asked to facilitate one small group activity a week (10-15 minutes) and four whole-group activities a week (5-15 minutes each activity). In addition, children independently participated in a computer game twice a week (10 minutes each session) and parents were sent weekly letters with suggested activities. Teachers participated in a total of 6 days of training, monthly 2-hour meetings, and once a month coaching sessions.

The researchers used a randomized controlled trial to evaluate the effectiveness of Building Blocks for children who were age-eligible for kindergarten the year following the study. Children completed the Early Mathematics Assessment (a researcher-designed assessment developed years prior in earlier studies of the curriculum) before and after the intervention. Their performance was evaluated relative to a comparison group of children who participated in an alternative curriculum (Pre-K Mathematics, described in a subsequent section) and a control group who participated in the standard classroom curriculum. The authors used Hierarchical Linear Modeling (HLM) to model differences in outcomes due to nesting of children within classrooms. The children in the Building Blocks classrooms performed significantly better than children in the comparison classrooms (d = .47) and the control classrooms (d = 1.07). In addition, the children in the comparison classrooms also significantly outperformed the children in the control classrooms (d = .64).

Measurement-based Head Start mathematics curriculum (Sophian, 2004). One of the primary goals of Sophian's intervention was to support 3- and 4-year-olds' understanding of units; specifically, that one can quantify something in different ways depending on the unit chosen (e.g., number, height, weight, volume). Another goal was to help preschoolers understand that quantities can be composed (i.e., two or more things added together to make one) and decomposed (i.e., one thing can be separated into two or more parts). As such, the intervention placed a heavy emphasis on measurement and geometry as ways to teach about numbers. The intervention was conducted in three Head Start centers. Six Head Start centers served as comparisons/controls: three implemented a parallel intervention for literacy, but not math, and three conducted business as usual. The intervention centered on a weekly project that was conducted in small groups. These projects consisted of games, puzzles, and manipulatives that targeted number, measurement, or geometric concepts. Researchers provided the teachers with lesson plans and suggestions on ways to incorporate the theme of that week into everyday activities. In addition, suggested activities, based on the weekly theme, were sent home to parents, and teachers kept logs of the home activities children completed.

Children completed the math subscale of the Developing Skills Checklist (DSC; 1990) and a researcher-designed assessment of measurement and geometry skills. The analyses revealed a significant treatment effect holding age and pretest scores constant, $F(2, 118) = 6.00$, $p < .01$; $\eta_p^2 = .09$ on the DSC and on the supplemental measure, $F(2, 118) = 5.33$, $p < .01$; $\eta_p^2 = .08$. The researchers observed that the children who participated in the intervention scored significantly higher than children in the other two conditions. The researchers also found that older children outperformed younger children on DSC post-test scores, $F(1, 118) = 5.13$, $p < .05$.

Pre-K Mathematics and Home Intervention (Starkey et al., 2004). The intervention program in this study consisted of a classroom component and a home component. The classroom component entailed a broad curriculum, *Pre-K Mathematics*, which was designed to aid the development of 4-year-olds students' informal mathematical knowledge. It included activities closely aligned with the National Council of Teachers of Mathematics (NCTM; 2000) pre-k to Grade 2 standards. The units included: enumeration and number sense; arithmetic reasoning; spatial sense; geometric reasoning; pattern sense and unit construction; non-standard measurement; and logical relations. Teachers conducted the activities in small groups of four to six children. Each activity lasted 20-30 minutes and occurred twice a week. Computer software with math activities was provided to each classroom, and teachers were encouraged to create or supplement a mathematics learning center within the classroom. The home component was designed to supplement the classroom activities. Parents were given training and materials for conducting activities at home. Teachers participated in a training program that consisted of a 5-day workshop in the summer before the start of school and 4-day workshop in the winter. In addition, teachers received on-site training from a research assistant once a month for the duration of the intervention.

The researcher's utilized a successive cohort design. Children enrolled in the classrooms during the first year of the project did not receive the intervention children while children enrolled in the second year of the study received the intervention. Children completed the Child Math Assessment, a measure designed by the researchers specifically for this study. Children who received the intervention scored significantly higher on the post-test measure, $F(1, 76) = 267.89$, $p < .001$. Children who received the intervention also scored significantly higher on the post-test than children who did not participate in the intervention, $F(1, 158) = 26.12$, $p < .001$. The researchers observed that children from lower SES families made greater gains ($d = 2.17$) than children from middle SES families ($d = 1.52$). However, when the researchers compared the test scores of low SES and middle SES children, they found that low SES children who participated in the intervention performed similarly to middle SES children who did not receive the intervention.

Summary. The long-term interventions generally were similar in scope, although Sophian's (2004) curriculum had a stronger emphasis in measurement and geometry, and Clements et al. (2008) and Starkey et al. (2004) supplemented their teacher-led activities with independent practice using computer software. While medium to large effects were found in each evaluation, only one research team (Clements et al.) evaluated the impact of their intervention compared to the effects another research-based curriculum (Pre-K Mathematics; Starkey et al.). As with the short-term intervention studies, reliability of the measures used in the long-term studies is a concern when interpreting the results of two of the three long-term interventions. Cronbach's alphas of .78 and .37 were reported for the measures used in Sophian et al.'s study. Starkey et al. did not report reliability information on their self-developed assessment. Another limitation of the Starkey et al. study is that children in the comparison condition were not assessed at pretest. The authors failed to provide information regarding the sampling method used, so it cannot be assumed the initial skill levels of intervention and comparison children were similar.

Conclusions and Rationale for the Current Study

Developing sound mathematical skills is an important task for all humans. Strong math skills have been correlated to wages after high school graduation and college completion (Dougherty, 2003; Murnane, Willett, & Levy, 1995). Recent evidence suggests a link between number sense and early math skills in kindergarten students and subsequent performance on standardized mathematical tests many years later, even into high school (Duncan et al., 2007; Halberda et al., 2008). Despite evidence that early numeracy skills continue to receive less attention than early literacy skills (Graham et al., 1997).

The development of early numeracy skills is complex and highly variable (Ginsburg et al., 2006). While many preschool-aged children demonstrate an impressive array of early mathematical skills (e.g., counting, quantity comparisons, estimation, simple addition/subtraction), not all children develop a strong foundation of early mathematical skills before entering school. One group historically at-risk for experiencing mathematical difficulties is children living in poverty (Aunola et al., 2004; Jordan et al, 2006; Lee & Burkam, 2002). Young children living in poverty enter school with fewer early numeracy skills than their middle income peers (Aunola et al., 2004; Jordan et al., 2006; Klein et al., 2004).

Several research teams have developed and evaluated the effectiveness of early numeracy interventions. While most of these interventions have demonstrated medium to large effects, there are a few limitations to these studies. For instance, many short-term interventions require a "specialist" to facilitate the intervention with children outside of the classroom. Preschool centers may not have the resources available to hire and train such a specialist. A second limitation of many of the interventions is the narrow focus of the skills targeted. Not all of interventions targeted the five domains (i.e., number and operations, geometry, measurement, data analysis, and algebra) identified by the NCTM (2000). A final limitation of the majority of early mathematics interventions is the fact the researcher-designed assessment of mathematics skills were used to measure the impact of the interventions, instead of norm-referenced, commercially available measures, which may introduce bias into the results.

The purpose of this study was to evaluate a short-term early numeracy intervention designed specifically for use in Head Start classrooms. Several factors differentiate this study from many of the previous early mathematics intervention studies. First, the intervention activities are aligned with three sets of standards: the Head Start Outcomes Framework (ACF, 2000; see Appendix A), NCTM Curriculum Focal Points (2006; see Appendix B), and the Pennsylvania Learning Standards for Early Childhood (2007; see Appendix C). Second, the scope and sequence of intervention activities is based on research conducted by Baroody et al. (2006), Clements (2004), and others. Third, multiple measures were used to assess the impact of the intervention on various skill domains. Finally, information about the mathematics instruction occurring in the comparison classrooms was recorded to aid in interpretation of the outcomes.

Research Questions

The purpose of this study was to determine if Number*Fun* promotes significant growth in Head Start children's early numeracy skills. Three research questions were addressed in the current study. The first, and most important, question is whether Number*Fun* has an impact on the development of early numeracy skills in Head Start children. The second and third questions provide a context for interpreting the results of the first question. That is, do Head Start intervention teachers implement the Number*Fun* program with fidelity, and is the implementation of the Number*Fun* early numeracy intervention regarded as feasible and practical by Head Start teachers?

Chapter 3

Method

Participants

Participants for this study were recruited from Lebanon County Head Start (LCHS) classrooms. The teacher recruitment letter and consent form are in Appendices D and E. Ten Head Start teachers expressed interest in participating in the project. One teacher declined upon follow-up. The remaining teachers were divided into two groups based on the age of the children in their classrooms. Six classrooms served both 3- and 4-year-olds, while three classrooms served only 3-year-olds. The classrooms in each group were then randomly assigned to intervention condition. Of the 3- and 4-year-old classrooms, two of the classrooms were located in the same building so they were treated as a unit to avoid treatment contamination. The remaining 3-year-old classroom was not selected to participate in the study. After the teachers were notified of their assignments, an intervention teacher informed the investigator that the time commitment was too great, and she was reassigned to the comparison condition. A randomly selected comparison teacher was assigned as a replacement.

Prior to the start of the study, one comparison classroom teacher was promoted to a non-teaching position and withdrew from the study. A second comparison teacher resigned from her position after the first wave of data collection. The investigator was unable to a find replacement for either teacher. After the second wave of data collection, but before the start of the intervention, an intervention teacher who participated in a pilot study of the Number *Fun* intervention resigned from her position. The substitute teacher who took over the classroom declined to participate in the study but allowed research assistants to assess participating children during the third wave of data collection. All participating teachers were English speaking Caucasian females. Additional teacher characteristics are displayed in Table 2.

Table 2. *Teacher Demographic Information*

	Degree	Certification	Years Teaching Experience: Total	Years Teaching Experience: Pre-K	Hours Math PD
Intervention					
Teacher 1	Bachelor	Elementary	9	4	0
Teacher 2	Associate	None	12	12	0
Teacher 3	Master	Secondary	7	3	0
Comparison					
Teacher 4	Master	Elementary	16	11	15
Teacher 5	Bachelor	Early Childhood Elementary	8	6	0
Teacher 6[a]	Bachelor	Early Childhood Elementary	21	21	0

Note. Math PD = professional development in mathematics.
[a]Teacher 6 withdrew from the study after the second wave of data collection, before the start of the intervention. Children in her classroom continued to be assessed, but the new teacher declined to participate in the study.

Children were recruited to participate in the study using an implied consent procedure. At the beginning of the academic year, teachers distributed consent forms and letters explaining the purpose of the study to families of children enrolled in their classrooms. Papers also were sent home to families of children who enrolled in their classrooms after the start of school year but before the start of the intervention. Both the letter and the consent form were written in English and Spanish at an appropriate reading level (see Appendices F to I). Parents and guardians were asked to return the signed consent form in a self-addressed, postage paid envelope if they did *not* want their children to participate in project related testing. No parents returned the consent forms.

A total of 96 children completed at least one measure during the course of this study. At the end of the intervention, the participants ranged in age from 41 to 65 months, with a mean age of 56.31 months (*SD* = 6.54). Approximately half (49%) of the children were male. Child demographic characteristics are presented in Table 3.

Table 3. *Demographic Characteristics of Participants by Classroom (n=96)*

	Intervention			Comparison		
	1 (*n* = 12)	2 (*n* = 17)	3 (*n* = 18)	4 (*n* = 18)	5 (*n* = 15)	6 (*n* = 16)
Age						
3-year-old	100%	88%	33%	17%	100%	37%
4-year-old	---	12%	67%	83%	---	63%
Gender						
Male	33%	53%	39%	72%	53%	37%
Female	67%	47%	61%	28%	47%	63%
Race/Ethnicity						
Caucasian	75%	77%	94%	100%	60%	94%
Hispanic	17%	17%	---	---	33%	---
Asian	---	---	---	---	---	6%
Biracial	8%	---	6%	---	7%	---
Other	---	6%	---	---	---	---
Primary Language						
English	92%	94%	100%	100%	87%	94%
Spanish	8%	6%	---	---	13%	---
Other	---	---	---	---	---	6%

Potential demographic differences between children in intervention versus comparison classrooms were examined. There were far fewer 4-year-olds (i.e., children who were age-eligible for kindergarten the following year) in the intervention condition (n = 14) compared to the comparison condition (n = 25), and age in years was found to be dependent on condition. However, when comparing intervention and comparison participants based on age in months, there is not a significant difference in age ($t(94) = 1.69$, $p = .09$). Gender, race/ethnicity, and primary language were not dependent upon condition. The results of the Chi-square Tests of Independence are reported in Table 4.

Table 4. *Comparison of Demographic Characteristics Between Participants in Intervention and Comparison Conditions*

	Intervention	Comparison	Tests of Independence		
	(n=47)	(n=49)	χ^2	*df*	*p*
Age					
3-year-old	70%	49%	4.48	1	.03
4-year-old	30%	51%			
Gender					
Male	43%	55%	1.51	1	.22
Female	57%	45%			
Race/Ethnicity					
Caucasian	83%	86%	.00[a]	1	.98
Hispanic	11%	10%			
Asian	---	2%			
Biracial	4%	2%			
Other	2%	---			
Primary Language					
English	96%	94%	.01[b]	1	.91
Spanish	4%	4%			
Other	---	2%			

Note. [a]Based on Caucasian and Hispanic participants only (n=95).
[b]Based on English and Spanish speaking participants only (n=91).

Number*Fun* Early Intervention Program. Number*Fun* is comprised of three units. Unit 1 focuses on skills such as counting words 1-10; numerical quantities 1-5, basic geometric shapes, measuring length using non-standard measures, and extending simple patterns (AB-AB). Unit 2 emphasizes counting words 11-20, numerical quantities 1-10, spatial sense, measurement concepts, extending more complex patterns (e.g., ABBC-ABBC), and interpreting graphs. Some of the skills focused on in Unit 3 include counting words 21-30, numerical quantities 11-20, adding/subtracting small quantities, number line estimation, and measuring weight. Across units, 65% of activities fall within the number and operations domain, 22% fall within the patterns and measurement domains, and 13% fall within the geometry and spatial sense domains. These percentages approximate guidelines provided by NCTM. Appendix J provides description of the specific skills targeted. Examples of Number*Fun* activities are found in Appendices K to M.

Number*Fun* has three types of activities. Primary activities typically occur in small- or whole-group settings and last for approximately 10 to 20 minutes. Most activities use concrete objects for children to manipulate to help them practice the target skill [e.g., cutouts of everyday objects, picture cards (index cards with stickers adhered to them), number cards (index cards with written numerals), craft supplies (pompons, ribbons), and number lines]. Other activities incorporate books, songs, and life-size board games. Activities progress in difficulty over the course of the unit. For example, early in a unit, children match quantities using picture cards, then later match a quantity with a numerical symbol using picture and number cards, then later order a set of quantities from least to most. Many primary activities are repeated to create additional practice opportunities.

Transition activities are brief activities (5 minutes or less) designed to be conducted during periods where children are moving from one location or activity to the next, such as being dismissed from circle time to free play, or lining up to leave the classroom. The primary purpose of these activities is to reinforce skills taught during the primary lesson and provide additional opportunities for practice. For example, if the primary activity required children to sequencing written numerals 1-5 from least to most, a corresponding transition activity would be placing number cards 1, 2, 4, and 5 in order in front of the child and asking the child to identify the missing number before dismissing the child from circle-time to brush his teeth.

The last type of Number*Fun* activity is the math center activity. Math center activities are introduced to children during circle time and then made available to the children for independent exploration during free play. Because these activities are designed for independent exploration, they contain built in features providing immediate feedback to children. For instance, one activity requires children to determine the number of objects in a set (e.g., 4 puppies) and find the corresponding written numeral (e.g. 3, 4, 5) on the bottom of the card. The child makes her selection by placing her index finger in a hole beneath the number. The child then turns the card around. If the child sees a star above her finger, then she knows she answered correctly. If not, she knows she needs to try again. Each math center activity is made available to the children for one week.

Measures

Mathematics achievement. Two measures were used to assess early mathematics skills in Head Start children. The Test of Early Mathematics Ability, 3rd Edition (TEMA-3; Ginsburg & Baroody, 2003) and the Early Arithmetic and Learning Indicators (EARLI; DiPerna & Morgan, 2006) Numeracy probes are described below.

***TEMA-3 (Ginsburg et al., 2003)*.** The TEMA-3 is an individually administered, norm-referenced test of early mathematics skills that can be completed by children ages 3 - 9. It is available in two forms (A and B) to allow for pre- and post- testing. The TEMA consists of 72 items intended to sample informal (number skills, number-comparison facility, calculation skills, and understanding of concepts) and formal (numeral literacy, mastery of number facts, calculation skills, and understanding of concepts) mathematics skills. The

TEMA-3 was standardized on a sample of 1,228 children from four major geographic regions in the United States in the fall of 2000 and the spring of 2001. The norms were derived by weighting the sample according to participants' ethnicity, region, and gender and 1999 census data. Both forms of the TEMA-3 have demonstrated high internal consistency with coefficient alphas ranging from .92 to .96 across all six age groups. The test also proved to be sufficiently stable over a 2-week period with test-retest reliability ranging from .82 to .93. Validity was examined through item discrimination and item difficulty. The average item discrimination coefficient across the six age groups was .55 for Form A and .58 for Form B. Median item difficulties ranged from .15 to .87 across forms and age groups. The TEMA-3 was moderately correlated with select subtests from the KeyMath-R/NU and the Woodcock Johnson-III Test of Achievement (WJ-III ACH). The TEMA-3 was moderately to highly correlated with subtests from the Diagnostic Achievement Battery-3rd Edition (*r*s = .65 to .84) and the Young Children's Achievement Test (*r* = .91). For the pre-test and post-test data from the current study, coefficient alphas were .94 and .96, respectively.

EARLI Numeracy Probes (DiPerna et al., 2006). The EARLI Numeracy Probes are a set of progress monitoring measures developed to assess early mathematics skills of preschool-aged children. The measure consists of six developmentally appropriate probes designed to be administered in the beginning, middle, and end of each school year. Each probe consists of six forms, one form per age (i.e., 3- or 4-year-olds) and one per time point, varying in item content and difficulty. The forms are linked by four to eight items that appear on all forms. In order to facilitate comparison of the different forms of each probe, the scores obtained from the participants were scaled and vertically equated using BILOG-MG (Grouping, in which a child can earn a score of 0, 1, or 2, was scaled and vertically equated using MULTILOG). The parameters used to scale the forms were based on the original EARLI field-test sample. The probes were created using Microsoft's PowerPoint program and are administered via a laptop computer. Typical administration time for the six numeracy probes is approximately 12 minutes for both 3- and 4-year-olds. A description of each probe is provided below.

Counting Aloud was designed to measure counting ability by asking examinees to begin with the number 1 and orally count as high as they can for 1 minute. Examinees receive 1 point for each number recited in the correct sequence until the first mistake made, for a total possible score of 60.

Counting Objects measures the ability to count using one-to-one correspondence by asking examinees to count a set of objects on the laptop screen as quickly and accurately as they can. Each form consists of eight items, four of which are present on all forms. Examinees receive 1 point for each test item answered correctly within the 10-second time limit, for a possible score of 8. Total administration time for this probe is 1 minute 20 seconds.

Grouping assesses the ability to subitize objects in a small set quickly and accurately. Each form consists of 11 items, six of which appear on all forms. For each item, the examinees are presented with two squares that contain one or more black dots and asked to quickly tell the examiner how many dots are in each box. Examinees receive 1 point for correctly identifying the number of circles in each box within the 10-second time limit, up to 2 points per item, for a possible total score of 22. Total administration time for this probe is 1 minute 50 seconds.

Measurement is designed to assess knowledge of measurement concepts (e.g., *bigger, smaller, more, less*). Each form consists of 12 items, eight of which appear on all forms. Examinees receive 1 point for each test item answered correctly within the 10-second time limit, for a possible total score of 12. Total administration time for this probe is 2 minutes.

Numbers & Shapes assesses number and basic geometric recognition skills by requiring examinees to name numbers and shapes presented in isolation as quickly and accurately as they can. Each form consists of 12 items, six of which appear on all forms. Examinees receive 1 point for each test item answered correctly within the 5-second time limit, for a possible score of 12. Total administration time for this probe is 1 minute.

Pattern Recognition assesses an examinee's skill in recognizing and completing simple patterns using combinations of three shapes (i.e., circle, square, and triangle). Each form consists of 10 items, five of which appear on all forms. Examinees are presented with an incomplete pattern of shapes (e.g., [blank space], circle, triangle, triangle) and asked to name which shape finishes the pattern. Examinees receive 1 point for each test item answered correctly within the 10-second time limit, for a possible score of 10. Total administration time for this probe is 1 minute 40 seconds.

The EARLI Numeracy Probes were piloted on a sample of approximately 250 children enrolled in Head Start. Reliability and validity evidence for the EARLI Numeracy probes was within acceptable ranges (DiPerna, Morgan, Lei, Reid, & Wu, 2006; Reid, Morgan, DiPerna, & Lei, 2006). Cronbach's alphas across the three time points for the all of the probes ranged from .61 to .97 for the 3-year-olds and .69 to .98 for the 4-year-olds, with the majority of alphas falling above .80 (DiPerna, Morgan, & Lei, 2009). Correlations between the EARLI Numeracy Probes and the Applied Problems and Quantitative Concepts subtests from the Woodcock-Johnson III (WJ III) Tests of Achievement ranged from .-.26 to .69 for 3-year-olds and .09 to .79 for 4-year-olds. Internal consistency and item analysis statistics for the current sample (see Appendix N) are similar to those reported by DiPerna et al. (2009).

Teacher information. Teachers provided information about their demographic characteristics and educational background (i.e., degree, certification, years experience, professional development experiences). In addition, they provided information about their classrooms (i.e., time spent doing literacy, science, mathematics, and social skills building activities).

Classroom instruction and intervention implementation fidelity. Two methods were used to gather information regarding the type and quality of instruction that occurred in intervention and comparison classrooms: a classroom observation protocol and activities logs. These instruments also were used to assess the degree to which teachers implemented Number*Fun* with fidelity.

Classroom Observation of Early Mathematics Instruction (COEMI). To assess the quality of mathematics instruction, all of the teachers were observed as they facilitated math activities in their classrooms using a researcher-made observation protocol. The protocol consisted of 22 items, on a 5-point Likert scale (0 = *Not observed*; 1 = *Strongly disagree*; 4 = *strongly agree*). The items were divided into four categories: Organization; Engagement; Quality Instruction, and Promotion of Mathematical Thinking. Organization items were designed to assess how well the teacher was prepared for an activity and managed activity materials. Engagement items focus on the teacher's attempts to engage and actively involve children in the activity. Items under Quality Instruction assess a variety of factors that promote learning (e.g., having high expectations, providing several opportunities for practice). Promotion of Mathematical Thinking items focus on how well the teacher encourages children to make informed guesses and evaluate their own and others responses. Four additional items were used to assess fidelity of implementation for Number*Fun* activities. Observations occurred during regularly scheduled mathematics activities and lasted approximately 45 to 60 minutes. Trained observers rated teachers on the 22 items immediately after observing a mathematics activity and recording notes. A separate form was completed for each activity observed. A copy of this measure can be found in Appendix O.

Early mathematics daily activity log. All comparison classroom teachers were provided with a log to record the math activities they facilitated during each school day. They were asked to provide information regarding the duration of the activity in minutes, to describe the activity briefly, and to record the learning objectivities. Completion of the log was monitored on a monthly basis. For an example of the log, please refer to Appendix P.

NumberFun daily activity log. The intervention teachers were provided with a log to record their daily math activities. For Number*Fun* activities, the teachers were asked to record the number corresponding to the activity in the intervention manual, the duration of the activity, and if activity extensions provided in the manual were utilized (*none, downward, upward*). Additionally, the teachers were asked to rate the Number*Fun* activities in terms of ease of implementation (*easy, moderate, difficulty*), difficulty level (*too easy, age-appropriate, difficult*), and level of child engagement (*not at all, a little, very much*). Finally, the teachers were asked to record any modifications made to the activities not suggested in the manual. The teachers also were asked to record information about any non-Number*Fun* mathematics activities they facilitated (i.e., duration, brief description, and learning objectives). Completion of the log was monitored on a monthly basis. For an example of the log, please refer to Appendix Q.

Number*Fun* intervention teacher satisfaction. At the conclusion of the study, teachers were asked to rate the components of Number*Fun* (e.g., manual, materials, activities, implementation) and their overall satisfaction with the program. The form consisted of 25 Likert items (1 = *strongly disagree*; 4 = *strongly agree*) and four open-ended questions that required written responses. A copy of the form can be found in Appendix R.

Procedures

Training of research assistants. Research assistants were recruited from the school psychology program at the Pennsylvania State University and from the Lebanon County area. All of the research assistants had prior data collection experience, and 2 out of 4 had previous experience assessing preschool-aged children. All research assistants completed the Office of Research Protections certification and had current criminal record clearances. Research assistants were compensated for their time.

***Child outcome measures*.** Prior to data collection, research assistants attended a 4-hour training to review and practice administering the measures (TEMA-3 and EARLI Numeracy Probes). To demonstrate mastery of the assessments, the researcher used a checklist (see Appendix S) to rate the research assistant's administration proficiency. Corrective feedback was provided, and the research assistants continued to practice administration until 100% accuracy was obtained. All data collectors obtained mastery. Research assistants were observed administering each measure in the field at least once. All research assistants obtained 100% accuracy during the field observation.

***Classroom observations*.** Research assistants attended a 2-hour training to review and discuss the components of the COEMI. A codebook for the COEMI was distributed, and a 3-hour practice observation in a local preschool classroom was conducted to establish inter-rater reliability. Practice observations continued until the investigator and research assistant established at least 85% agreement on two consecutive observations. To maintain consistency in ratings across the duration of the intervention, inter-rater reliability was assessed on 10% of observations ($\kappa = .66$, $p < .001$).

Training of teachers. Training for the intervention classroom teachers consisted of a half-day workshop conducted before the intervention. The first part of the training focused on the empirical basis for the program, as well as the primary objectives, purpose, and rationale. Next, the teachers were acquainted with the content and organization of the Number*Fun* manual and materials. The last portion of the training focused on providing the teachers with strategies to facilitate discussion about early numeracy concepts with children in their classrooms. In addition to this training, classroom teachers were provided with monthly individual coaching (support and feedback) in the classroom regarding their implementation of the program.

Data collection. Participants' early numeracy skills were assessed prior to the intervention and after the conclusion of the intervention. Children were asked for their assent to participate in the testing. If the child declined initially, the research assistant moved to the next child on the list; then, approached the child again at a later time. If the child declined the second request they were not tested. During each data collection wave, children completed the TEMA and EARLI Numeracy Probes in separate test sessions within a 2- to 4-week period. The measures were presented in a counter-balanced order. The average time in weeks between the first and second waves was 16.22 (range = 15.14 to 18.00). Children received prizes (small trinkets and stickers) for completing the assessments.

Research staff conducted classroom observations using the COEMI in the fall, prior to implementation, and once a month during the course of the intervention. Teachers provided research staff with their daily schedules and offered days when researchers could visit the classroom to view the teacher conducting math activities. Due to security issues at the various sites, researchers were required to give advance notice of their classroom visits. Observations lasted between 45 and 60 minutes, depending on the number and duration of mathematics activities facilitated during that period.

Number*Fun* implementation. Intervention teachers were provided with an intervention manual that contained the schedule of the Number*Fun* activities they were to facilitate on a daily basis. Teachers were asked to facilitate three activities a day: a 10 - 15 minute teacher-led small- or large-group activity; a brief teacher-led activity that occurred during a transition period; and a child-led activity made available during free play. All of the materials needed to facilitate the activities were provided in advance by the investigator. Teachers were instructed to record information about each facilitated activity in their Number*Fun* Intervention Daily Activity Log described above. Any non-Number*Fun* activities, such as calendar math, were also recorded in the log. Implementation of the activities was observed on a monthly basis. Based on the observations, teachers were provided with feedback and suggestions as needed. Teachers were compensated for the time they spent working on Number*Fun* activities.

Comparison classroom procedures. Comparison teachers attended a 30-minute meeting where they were briefly introduced to the purpose of the project and the data collection procedures. The teachers were instructed to proceed with their usual practices in the classroom and to record information about the activities they facilitated in the Early Mathematics Daily Activity Log. The curriculum in all of the classrooms was Creative Curriculum, which is described by its publisher as being “Nationally known for its forward-thinking, comprehensive, and rigorously researched mode… that supports active learning and promotes children's progress in all developmental areas” (Teaching Strategies, Inc., 2010). No additional mathematics specific curriculum was used in any of the comparison classrooms. The comparison teachers were observed on a monthly basis; however, no feedback or suggestions were offered. Teachers were compensated for the time they spent in Number*Fun* data collection related activities.

Design and Analysis

The primary aim of this quasi-experimental study was to determine if children who participated in the Number*Fun* intervention demonstrated significant growth in early numeracy skills above that of children who did not receive the intervention, but participated in the standard curriculum. In order to provide a context for interpretation of the child outcomes, descriptive statistics were run on measures of classroom instruction, intervention implementation fidelity, and teacher acceptability of the intervention.

To test the hypothesis that intervention condition would predict post-test scores after controlling for pretest scores and demographic characteristics, hierarchical regression analyses were run for each early mathematics outcome measure. Hierarchical regression was selected for the analyses because the primary interest of this study was the proportion of R^2 explained by the treatment condition after the variance of a set of covariates was partitioned.

The equation tested was as follows:

$$MATH = b_0 + b_1(PRE) + b_2(AGE) + b_3(RACE) + b_4(SEX) + b_5(GROUP) + e.$$

Where MATH is post-test scores for any one of the outcome measures; PRE is the corresponding pretest scores on a mathematics measure; AGE is age in months grand mean centered; RACE is a dummy coded variable (0 = Caucasian, 1 = other); Sex is a dummy coded variable (0 = male, 1 = female); GROUP is a dummy coded variable representing the treatment condition (0 = comparison, 1 = intervention), and *e* is the error term.

Pretest scores were entered into the first step of the analyses to account for differences in early mathematics skills among children prior to the intervention. The set of demographic characteristics were entered into the second step. Due to the variability in age composition of the conditions, age in months was included in the model. It was anticipated that older children, who were likely to be more cognitively advanced and have had completed an additional year of Head Start, might respond differently to the intervention. Race and gender were entered into the prediction equation because previous studies of children in preschool and kindergarten have found gender and ethnic differences in mathematics achievement (e.g., Arnold et al., 2002; Jordan, et al., 2006). The demographic variables were entered in the same step because their individual contribution to R^2 was not the focus of this study. The intervention condition variable was entered into the last step. Regression coefficients with a p-value of .05 or less were considered statistically significant.

As recommended by Cohen, Cohen, West, & Aiken (2003), Pedhazur (1997), and Tabachnick and Fidell (2001) several assumptions were tested to determine if the data collected were appropriate for analyses. First, data were examined for outliers and influential cases. Second, the assumption of normality was explored using histograms and normal P-P plots. Next, linearity and homoscedasticity were assessed by plotting standardized residual against standardized predicted scores and examining partial residual plots. Finally, independence or errors was investigated by obtaining Durbin-Watson statistics and calculating the intraclass correlation.

When interpreting the results of the regression analyses, it is important to recognize the nested nature of the data. The data analyzed in this paper has a three-level hierarchical structure. Measurement time points are nested within individual children and individual children are nested within classrooms. Due to the child-classroom nesting, children's performance on math measures may be more correlated to the performance of children in their own classrooms than to children in other classrooms because they share similar environments and experiences (e.g., instruction by the same teacher, opportunity to play with the same educational materials). These correlations, which result in systematic variation in error terms, are potentially problematic because they result in a violation of the assumption of independence of observation required by most statistical procedures (Baldwin, Murray, & Shadish, 2005; Hoyle, Georgesen, & Webster, 2001; Kenny & Judd, 1986; Kenny & La Voie, 1985). While the estimates produced by the analysis will not be affected by non-independence, the estimation of error variance may be biased which leads to unreliable standard errors and significant tests (Dorman, 2008; Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). As a result, there is increased risk of making a Type I error or a Type II error with a reduction in power (Kenny et al, 2002).

While researchers and statisticians have been aware of the complications of nested data for several decades, historically, education researchers either completely ignored the nested factor while analyzing individual differences or they analyzed group means while ignoring individual information. Neither method is ideal. Deciding to ignore the nested data and conduct analyses based on individuals may result in biased standard error estimates and increased Type I errors (Kenny et al, 1985). Ignoring information about individuals and analyzing group means, however, may lead to increased Type II errors and underestimation of effect sizes (Siemer & Joorman, 2003). More recently, education researchers have employed techniques such as hierarchical linear modeling (HLM) which makes use of information provided by both individuals and groups. While HLM is the preferable approach when data is nested, it requires a larger sample size, particularly a large number of groups (e.g., classrooms).

While conducting an HLM analysis is not an option with small sample sizes or limited number of groups, other options can be utilized. One method described by Hedges (2007) is to correct the *t* statistic obtained from analyses of individual differences that ignore the group factor. This correction is applied after the fact instead of accounting for nesting in the research design or in the original statistical analyses. Another approach is to account for the variance attributed to the group factor by including it in the original model and testing a group by teacher interaction (Siemer et al, 2003; Wampold & Serlin, 2000). Due to issues regarding small sample size and missing data, neither of these approaches is used in the current study. The level of analysis is the individual, and therefore the results should be interpreted with caution.

Treatment of Missing Data in Outcome Measures

One frequently encountered difficulty with longitudinal designs is missing data. Participants may withdraw from the study, they may not be available for all data collection waves, or they may not complete all of the items of an assessment. While missing data in longitudinal research is practically impossible to avoid, there are several methods available to handling missing data, some better than others (for comprehensive reviews on the treatment of missing data, see Allison, 2002; Graham, 2009; Schafer & Graham, 2002; Rubin, 1987). Methods such as listwise or pairwise deletion, where cases with missing data are deleted from the data file or from the analysis, or mean substitution of missing values, often lead to biased parameter estimates (Schafer et al., 2002). A better approach (Jeličić, Phelps, & Lerner, 2009) to treating missing data is to use multiple imputation, which replaces missing values with plausible values for each case based on information available from the model and accounts for the relationship between the variables based on the entire sample (Schafer et al., 2002). The result is a user-specified number of imputed data sets that maintain the means, variances, and correlations of the original data (Schafer et al., 2002).

Patterns of missingness can be described in one of three ways (Rubin, 1976). Missing completely at random (MCAR) indicates that a missing value on variable A is not related to the value of A itself (e.g., a child who has difficulty counting and would obtain a low score on a counting task is no more likely to have missing data for this task than children who would obtain a high score) and is not related to any other variables in the data set (e.g., younger children are no more likely to have missing data on a counting task than older children). Missing at random (MAR) is a less stringent condition where a missing value on variable A is not related to the value of A itself (as with MCAR); however, it may be related to an observed value on variable B (e.g., a missing value on the counting task is related to low-income status). Missing data also can be described as missing not at random (MNAR). MNAR indicates that the missing value on variable A is related to the value of A itself (e.g., not completing the counting task is related to low counting ability). When missing data is

MNAR, the missing data mechanism must be modeled when imputing values (Allison, 2002). When the conditions of MCAR or MAR are met, the missing data mechanism is considered to be ignorable.

In the current study, SPSS (PASW) 17.0 was used to ascertain the pattern of missingness for the outcome variables of the 96 cases in the data set. All of the outcome measures had some missing data, whereas 62% of cases had missing data. Approximately 28% of all outcome measure values were missing. Table 5 presents the percentage of missing values per outcome variable. The majority of missing data was due to participants being unavailable due to absence during a data collection wave; however, there were a few instances when participants refused to complete an assessment task. Based on this information, missing data are assumed to be MAR. Therefore, the missing data mechanism did not have to be modeled in the imputation model.

Table 5. *Distribution of Missing Values in Child Mathematics Measures (n = 96)*

Measure	Total Missing (%)	Reason for Missing Values		
		Attrition (%)	Unavailable (%)	Refusal (%)
Pretest				
TEMA	34 (35.4)	9 (9.38)	25 (26.04)	---
Counting Aloud	35 (36.5)	9 (9.38)	25 (26.04)	---
Counting Objects	35 (36.5)	9 (9.38)	25 (26.04)	1 (1.04)
Grouping	35 (36.5)	9 (9.38)	25 (26.04)	1 (1.04)
Measurement	36 (37.5)	9 (9.38)	25 (26.04)	2 (2.08)
Number Naming	34 (35.4)	9 (9.38)	25 (26.04)	---
Pattern Recognition	36 (37.5)	9 (9.38)	25 (26.04)	2 (2.08)
Post-test				
TEMA	25 (26.0)	11 (11.46)	13 (13.54)	1 (1.04)
Counting Aloud	31 (15.6)	11 (11.46)	16 (16.67)	4 (4.17)
Counting Objects	29 (14.6)	11 (11.46)	16 (16.67)	2 (2.08)
Grouping	31 (15.6)	11 (11.46)	16 (16.67)	4 (4.17)
Measurement	29 (14.6)	11 (11.46)	16 (16.67)	2 (2.08)
Number Naming	29 (14.6)	11 (11.46)	16 (16.67)	2 (2.08)
Patter Recognition	30 (15.6)	11 (11.46)	16 (16.67)	3 (3.13)

SPSS Missing Values 17.0 was used to conduct multiple imputation. The method of imputation selected was the fully conditional specification, which is an iterative Markov chain Monte Carlo (MCMC). The program makes random draws from a normal distribution of plausible values with the same mean and standard deviation as the original, incomplete data. When the user-defined maximum number of iterations is reached (20 for this study), that value is then imputed in the data set. This is repeated for each variable with missing data specified in the model. The number of imputed datasets (*m*) also is dependent on user specifications. Rubin (1976) wrote that *m* = 2 - 3 imputations should be sufficient in most cases; whereas Shafer and Olson (1998) recommended m = 3 - 5 and Shafer et al. (2002) recommended m = 20 imputations. More recently, Graham, Olchowski, and Gilreath (2007) suggested that even more imputations might be needed due to their findings that as m decreases, MSE and SE increases and power decreases. In order to determine the appropriate m, it is important to know the fraction of missing information for each variable. Allison (2002) described the fraction of missing information as “an estimate of how much information about each coefficient is lost because of missing data” (p. 48) that takes into account the amount of missing values in a variable as well as the amount of missing values in related variables. The formula to calculate the fraction of missing information is shown below (Allison, 2002; Schafer et al., 1998).

$$\lambda = [r + 2 / (df + 3)] / (r + 1)$$

where the relative increase in variance due to missing data is defined as

$$r = [(1 + m^{-1})\ B / U$$

where

B = variance of the coefficients between regressions.

U = average of the squared, within-regression standard errors.

with

$df = (m - 1)\ (1 + r^{-1})^2$

The fraction of missing information is an estimate, and the values change for each imputed data set (Allison, 2002). Bodner (2008) recommended a simpler, yet conservative, method of estimating the fraction of missing information that controls for imputation variability and is particularly useful for multivariate missing data:

$$\lambda_{(hat)L} = 1 - n_L / n$$

where

n_L = number of cases available for the estimation of a given parameter using listwise deletion.

n = total sample size.

Bodner provided a table (p. 668) to determine the minimum number of imputations needed given $\gamma_{(hat)L}$. Given the current study, $\lambda_{(hat)L} = 1 - (60 / 96) = .375$. The value of .38 fell between values of .30 and .50 (95th percentile) in the table; therefore, Bodner recommends $m = 24 - 59$. Linear interpolation was used to determine the exact number of imputations needed.

$$m = m_0 + (\lambda_{(hat)L} - \lambda_0) * [(m_1 - m_0) / (\lambda_1 - \lambda_0)]$$

where

m_0 = number of imputations for lower value of λ in table.

m_1 = number of imputations for higher value of λ in table.

λ_0 = lower value of λ in table.

λ_1 = higher value of λ in table.

For the data in this study, $m = 24 + (.38 - .30) * [(59 - 24) / (.50 - .30)] = 38$. Therefore, 38 imputations were specified in SPSS.

A multivariate imputation model was then specified. The following variables were included in the model as predictors: treatment condition (0 = comparison; 1 = intervention), teacher (dummy coded), which form of the EARLI Numeracy probes would be administered (0 = 3-year-old; 1 = 4-year-old), and age in months at beginning of school year. Three waves of the mathematics measures were entered into the model, in the sequence in which they were administered, to be imputed and to serve as predictors for the other imputed variables.

The model type selected for the scale variables was linear regression and two-way interactions among categorical predictor variables were included. Constraints placed on the imputed values limited them to be within three standard deviations of the variable mean. The maximum number of iterations was set at 20, the maximum case draws was set to 100, and the maximum parameter draws was set to 20.

Once missing values were imputed, analyses were run on each imputed data set. The results from the imputed datasets were then pooled according to Rubin's (1987; Schafer et al., 2002) rules. Pooling the estimates is rather straightforward. The overall estimate is simply the average of the point estimates across imputed data sets,

$$\overline{Q} = \frac{1}{m}\sum_{j=1}^{m}\hat{Q}_j.$$

To obtain an overall standard error, one must calculate the within-imputation variance (the average of sampling variances),

$$\overline{U} = \frac{1}{m}\sum_{j=1}^{m}U_j\,.$$

and also calculate the between-imputation variance,

$$B = \frac{1}{m-1}\sum_{j=1}^{m}\left(\hat{Q}_j - \overline{Q}\right)^2.$$

The total variance is calculated as follows,

$$T = \overline{U} + \left(1 + \frac{1}{m}\right)B.$$

and the overall standard error is the square root of the total variance. Significance tests of the null hypothesis are calculated as follows,

$$t = \overline{Q}\Big/\sqrt{T}$$

with degrees of freedom,

$$df = (m-1)\left(1+\frac{m\,\overline{U}}{(m+1)B}\right)^2.$$

Because there are no published guidelines for creating pooled multiple correlations in multiply imputed datasets, ranges for R^2 and sr^2 are reported in the results section. Cohen's *d* values also were calculated using average point estimates via the following formula (Wolf, 1986):

$$d = 2r / (1 - r^2)^{-1}$$

To maximize comparability between the previous studies cited in the literature review and the current study, *d* effect sizes also were calculated between pretest and post-test scores for each condition and between the intervention and comparison groups' post test scores. These effect sizes were calculated via the following formula (Lipsey & Wilson, 2001):

$$d = (M_1 - M_2) / \{[(n_1\text{-}1) * \sigma_1^2] + [(n_2\text{-}1) * \sigma_2^2] / [(n_1 - 1) + (n_2 - 1)]\}^{-1}$$

Chapter 4

Results

Results of the current study are organized and reported in four sections. The first section reports data gathered to assess implementation fidelity, including classroom observations and activity logs. The second section provides descriptive data regarding the early mathematics instruction in comparison classrooms. The next section presents information regarding teacher acceptability of the Number*Fun* intervention. These first three sections provide a context for the fourth and final section, which reports results of the multiple regression analyses to test the effects of the Number*Fun* intervention.

Number*Fun* Implementation Fidelity

Direct observation data. Teachers were observed for 1 hour on three separate occasions during the course of the study (approximately once a month). During each observation, teachers facilitated between one and three mathematics activities. Of the activities observed, 83% were Number*Fun* activities. Activities, on average, lasted 9.33 minutes (*SD* = 4.45). The majority of activities (68%) occurred in a whole group setting, while 19% occurred in a small group setting and 13% occurred during transitions from one activity to another. As measured by the COEMI, teachers received overall ratings of fidelity in the *agree to strongly agree* range, with scores for individual items ranging from 2.5 to 4.0. Teacher 3 obtained the lowest overall fidelity score due to not following the activity schedule.

Teacher self-report. Of the 113 Number*Fun* activities, teachers reported facilitating an average of 95 activities. Teachers 1 and 2 conducted 108 (96%) and 107 (95%) of activities, respectively, while Teacher 3 conducted 71 (63%) activities. Overall, teachers followed the activity schedule provided in the manual 85% of the time; however, there was some variability among teachers. Specifically, Teachers 1 and 2 followed the intervention activity schedule 99% and 92% of the time respectively, while Teacher 3 followed the intervention activity schedule 72% of the time.

The Number*Fun* manual provided downward and upward extensions for most activities to accommodate the varying skill level of children participating in the intervention. The teachers reported utilizing these extensions for 7% of activities. In addition, teachers reported modifying 16% (range = 13% to 18%) of all activities in ways that were not suggested by the manual (described further in the Discussion section).

Characteristics of Early Mathematics Instruction

Information from classroom observations and teacher activity logs was compiled to provide a description of the type and quality of early mathematics instruction provided in the classrooms. Classroom observations occurred three times over the course of the intervention (approximately once a month) and lasted for approximately 1 hour, depending on the length and number of math activities facilitated.

Number*Fun* classrooms. At the conclusion of each observed activity, teacher instructional behavior was rated using the COEMI. Composite scores from the COEMI are presented in Table 6, while individual item responses are presented in Appendix S. Teachers received the highest scores in the Organization (M = 19.83, SD = .51) and Engagement (M = 14.44, SD = 1.65) domains, while scoring lowest in the Quality (M = 22.28, SD = 4.82) and Promotion of Mathematical Thinking (M = 15.56, SD = 3.90) domains. Examination of individual item scores revealed some consistencies across teachers. All teachers received lower average ratings for the following instructional practices: *provided several opportunities for practice*; *checked for understanding*; *corrected children's misunderstandings*; *elaborated children's ideas; encouraged reflection*; *encouraged evaluation of others' answers*; and *uses examples and non-examples*.

According to the Number*Fun* activity logs, teachers reported conducting 3.25 (*SD* = .35) mathematics activities a day for a daily total of 29.51 minutes (*SD* = 4.37) of mathematics instruction (Table 7). Numbers and operations skills were targeted in 84% of activities; geometry and spatial sense skills were targeted in 14% of activities; and patterns and measurement skills were targeted in 34% of activities (Note: Percentages do not total 100 because many activities target more than one skill.).

Comparison classrooms. During each observation, comparison teachers facilitated between two and three mathematics activities. Observed activities lasted between 3 and 20 minutes (*M* = 8:09, *SD* = 4:17). All mathematics activities occurred in a whole group setting, 55%, of which occurred during review of the calendar (e.g., counting days of the month, completing geometric patterns). In comparison classrooms, 84% of activities targeted numbers and operations (e.g., word counting, counting steps), 54% targeted patterns and measurement (e.g., completing patterns, graphing preferences), and 8% target geometry and spatial sense (e.g., filling in an oval with broken eggshells). Within the number and operations domain, teachers targeted counting (73%) and number recognition (18%).

Teacher instructional behavior during each activity was rated using the COEMI (see Table 6 for composite scores and Appendix T for item responses). Comparison teachers scored highest overall on the Organization subscale (*M* = 19.15, *SD* = .99) while scoring lowest on the Promotion of Mathematical Thinking subscale (*M* = 10.46, *SD* = 4.03). Within the Promotion of Mathematical thinking subscale, teachers received lower ratings on items such as: *encouraged reflection, elaborated children's ideas, uses examples and non-examples,* and *encouraged evaluation of others' answers*. Teachers were rated low on other items such as *provided several opportunities for practice* and *checked for understanding*.

Comparison classroom teachers recorded the math activities they facilitated throughout the project. On average, teachers reported conducting 2.25 activities (*SD* = .79) for 18.99 minutes per day (*SD* = 8.26). As far as the skill domains targeted by the activities, numbers and operations skills were targeted in 86% of activities; geometry and spatial sense skills were targeted in 10% of activities; and patterns and measurement skills were targeted in 52% of activities (Note: Percentages do not add up to 100 because many activities target more than one skill.).

Summary. Number*Fun* and comparison teachers received similar ratings (and their highest) on the organization domain of the COEMI (*M* = 19.83, *SD* = .51 and *M* = 19.15, *SD* = .99, respectively). For each of the remaining domains, the mean score of the intervention group was roughly 4 points higher than the comparison group mean. Teachers in both conditions received the lowest ratings in the promotion of mathematical thinking domain (*M* = 15.56, *SD* = 3.90 and *M* = 10.46, *SD* = 4.03 for the intervention and comparison conditions, respectively). Table 6 presents COEMI ratings for individual teachers across conditions.

Table 6. *Composite Scores for COEMI Domains and Fidelity Scale Averaged Across Observations*

	Number*Fun* Classrooms						Comparison Classrooms			
	1		2		3		4		5	
Subscales/Items	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Organization	19.71	.76	20.00	.00	19.83	.41	19.43	1.13	18.83	.75
Engagement	13.00	1.53	15.80	.45	15.00	1.10	11.29	2.50	10.67	1.63
Quality	18.29	5.56	24.00	1.73	25.50	1.38	20.29	3.99	16.83	3.76
Promotion of Math Thinking	14.29	4.46	17.00	4.42	15.83	2.79	12.71	3.45	7.83	3.06
Number*Fun* Fidelity	16.00[a]	.00	15.40	.55	13.17	1.83	---	---	---	---

Note. Scores based on 7 observations for Classrooms 1 and 4; 6 observations for Classrooms 3 and 5; and 5 observations for Classroom 2. Possible score range: Organization = 5 – 20, Engagement = 4 – 16, Quality = 7 – 28, Promotion of Math Thinking = 6 – 24, Number*Fun* Fidelity = 4 – 16. The teacher of Classroom 6 did not participate in this aspect of the study.
[a]Based on 3 observations.

Number*Fun* teachers reported conducting one more mathematics activity each day than comparison teachers. In addition, Number*Fun* teachers spent approximately 10.5 more minutes on early numeracy activities per day than comparison teachers. Table 7 presents information on the time spent in mathematics activities by teachers. The percentage of time spent in activities emphasizing number and operations was roughly the same across intervention and comparison conditions (84% and 86%, respectively). Comparison teachers spent a higher percentage of time on patterns and measurement activities than intervention teachers (52% compared to 34%).

Table 7. *Daily Average Number and Length of Mathematics Activities by Classroom*

	Number*Fun* Classrooms						Comparison Classrooms			
	1		2		3		4		5	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Number of Activities	2.95	.21	3.92	.43	2.88	.33	2.04	.60	2.47	.88
Activity Length[a]	10.06	.98	6.82	.79	11.26	2.13	7.86	2.29	8.69	2.29
Total Daily Minutes	29.62	3.01	26.52	3.11	32.38	7.07	16.61	7.57	21.29	8.30

Note. [a]In minutes.

Teacher Acceptability of Number*Fun* Intervention

Number*Fun* teachers completed a 25-item survey regarding their experience with Number*Fun*. Ratings ranged from 1 (*strongly disagree*) to 4 (*strongly agree*). Table 8 presents the composites scores for teachers' ratings within each domain. Appendix U presents the teacher ratings for each item on the survey. Average item ratings of satisfaction with the Number*Fun* manual ranged from 3.00 to 4.00, with an item related to the ease of following the manual receiving the lowest average rating. Average implementation satisfaction ratings ranged from 3.33 to 4.00. Average ratings for satisfaction with Number*Fun* activities ranged from 1.67 to 4.00. While most of the items in this domain received average ratings of 3 or higher, all the teachers endorsed *agree* or *strongly agree* when asked if the activities were too long.

Table 8. T*otal Teacher Satisfaction Ratings of NumberFun Intervention by Domain*

Domain	Teacher 1	Teacher 2	Teacher 3
Manual	17	19	17
Implementation	13	16	15
Activities	43	47	47
Satisfaction	12	12	12

Note. Possible score range: Manual = 5 – 20, Implementation = 4 – 16, Activities = 13 – 52, Satisfaction 3 – 12.

In addition to completing teacher satisfaction surveys, teachers provided information regarding characteristics of each activity in their daily activity logs. Specifically, teachers were asked to rate the difficulty of implementing each activity in the classroom, level of difficulty for the majority of children in the classroom, and the children's level of engagement while participating in the activity. Composite scores for each domain are provided in Table 9. Teacher 1's and 3's ratings indicated that they found the implementation of the activities to be easy to moderate and that the activities, on average, were age-appropriate in difficulty for most children. Teacher 1 rated the activities as moderately engaging for the children in her classroom; whereas Teacher 3 rated the activities as very engaging. There was little variability in Teacher 2's ratings. She rated the activities as easy to implement, age-appropriate in difficulty, and very engaging.

Table 9. *Teacher Ratings of NumberFun Activity Acceptability*

	Teacher 1 (n = 108)		Teacher 2 (n = 107)		Teacher 3 (n = 71)		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Ease of Implementation[a]	1.49	.63	1.01	.10	1.65	.51	1.34	.31
Difficulty Level[b]	2.14	.48	2.00	.00	2.04	.20	2.06	.22
Child Engagement Level[c]	2.16	.51	2.99	.10	2.83	.38	2.80	.27

Note. Possible score range: 1 – 3. [a]Lower scores indicate lower implement difficulty. [b]Lower scores indicate lower difficulty. [c]Lower scores indicate lower engagement.

Child Outcomes

To test the hypothesis that children participating in the Number*Fun* intervention performed significantly better than children in business-as-usual classrooms, separate hierarchical regression analyses were run for each criterion variable. Missing data were handled through multiple imputation procedures, which resulted in 38 imputed datasets.

Appendices V to AB present descriptive statistics for each imputed data set by probe. Pooled descriptive statistics of the TEMA-3 and EARLI Numeracy probes by curricular condition are presented in Table 10. Table 11 presents the distribution of early mathematics scores at pre- and post-test by classroom.

Table 10. *Pooled Means and Standard Errors of Early Mathematics Outcomes by Intervention Condition*

	Pretest				Post-test			
	Number*Fun* $n = 47$		Comparison $n = 49$		Number*Fun* $n = 47$		Comparison $n = 49$	
Measures	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
TEMA-3	7.88	0.95	7.76	0.96	10.63	1.29	13.50	1.53
Counting Aloud	-0.42	0.14	-0.64	0.09	-0.24	0.17	-0.30	0.18
Counting Objects	-0.94	0.27	-0.60	0.24	-0.48	0.17	-0.18	0.17
Grouping	0.73	0.16	0.62	0.17	1.29	0.19	1.29	0.19
Measurement	-0.72	0.17	-0.80	0.14	-0.48	0.19	-0.33	0.18
Number Naming	-0.60	0.11	-0.57	0.08	-0.41	0.14	-0.35	0.16
Pattern Recognition	-0.35	0.13	-0.36	0.12	-0.29	0.19	-0.24	0.18

Note. Raw scores are reported for the TEMA-3. IRT scaled scores are reported for Counting Aloud, Counting Objects, Grouping, Measurement, Number Naming, and Pattern Recognition.

Table 11. *Pooled Means and Standard Errors of Early Mathematics Outcomes by Classroom*

	Number*Fun* Classrooms						Comparison Classrooms					
	1		2		3		4		5		6	
Measures	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
							Pretest					
TEMA-3	7.65	1.04	4.54	1.29	11.18	1.64	9.32	1.89	4.42	1.02	9.14	1.63
Counting Aloud	-0.58	0.32	-0.49	0.21	-0.26	0.24	-0.61	0.18	-0.76	0.12	-0.56	0.14
Counting Objects	-0.78	0.48	-1.05	0.43	-0.93	0.49	0.04	0.42	-1.06	0.36	-0.87	0.36
Grouping	0.75	0.36	0.71	0.24	0.73	0.30	0.90	0.31	0.55	0.25	0.39	0.29
Measurement	-0.76	0.35	-0.73	0.26	-0.68	0.32	-0.38	0.19	-1.34	0.28	-0.77	0.23
Number Naming	-0.53	0.31	-0.79	0.12	-0.47	0.14	-0.45	0.14	-0.92	0.07	-0.37	0.16
Pattern Recognition	-0.34	0.30	-0.55	0.17	-0.15	0.22	0.09	0.20	-0.82	0.16	-0.44	0.14
							Post-test					
TEMA-3	7.34	1.88	9.67	1.96	13.73	2.20	16.29	2.82	8.38	1.68	15.17	2.70
Counting Aloud	-0.48	0.31	-0.29	0.25	-0.04	0.27	0.06	0.30	-0.75	0.23	-0.27	0.35
Counting Objects	-0.37	0.33	-0.61	0.27	-0.42	0.30	-0.07	0.30	-0.62	0.27	0.12	0.31
Grouping	0.75	0.33	1.15	0.24	1.76	0.31	1.45	0.32	0.61	0.29	1.74	0.32
Measurement	-0.68	0.37	-0.77	0.31	-0.08	0.31	0.01	0.26	-0.88	0.25	-0.21	0.37
Number Naming	-0.78	0.25	-0.66	0.20	0.08	0.26	-0.10	0.26	-0.70	0.19	-0.30	0.32
Pattern Recognition	-0.22	0.34	-0.46	0.29	-0.17	0.35	0.23	0.27	-0.62	0.24	-0.41	0.34

Note. Raw scores are reported for the TEMA-3. IRT scaled scores are reported for Counting Aloud, Counting Objects, Grouping, Measurement, Number Naming, and Pattern Recognition.

Relevant assumptions were tested to ensure data were appropriate for the proposed analyses. First, data were examined for outliers and influential cases. As recommended by Pedhazur (1997), all cases with standardized residuals greater than $|2|$ were further examined. No cases in any of the analyses exceeded criteria established for Cook's D (>1) or Mahalanobis Distance ($\chi^2(5) = 20.52$, $p = .001$). Linearity and homoscedasticity were assessed using two methods. First, standardized residuals were plotted against standardized predicted scores. Across analyses, residuals generally were scattered randomly and evenly around zero. Second, partial residual scatter plots were examined. Again, no evidence of severe nonlinearity or heteroscedasticity was observed. Normality was assessed by examining histograms and normal P-P plots of standardized residuals. No evidence of non-normality was observed across analyses.

Independence of residuals was assessed using two methods. First, serial dependency was tested using the Durbin-Watson statistic. All values were within acceptable ranges (1.00-3.00; Field, 2009). The degree of dependency related to participant clustering within classrooms was examined by calculating the one-way intraclass correlation[1] (ICC; McGraw & Wong, 1996). Table 12 presents the pooled ICC for each measure. When ICC equals zero, no dependence exists between participants' scores. In the current study, a small degree of dependence due to clustering is evident for all measures. However, Cohen et al., (2003) stated that "An ICC of .01 or .05 may seem very small; however, the actual alpha level of statistical tests increases dramatically even with such apparently small ICCs" (p. 538).

[1] One-way ICC formula for unequal group size (Cohen et al., 2003; Snijders & Bosker,1999):

$ICC = (MS_{treatment} - MS_{error}) / [MS_{treatment} + (\tilde{n} - 1)MS_{error}]$, where $\tilde{n} = M_n - [sd^2(n_j) / gM_n)]$ and where M_n = mean number of case per group, $sd^2(n_j)$ = variance of the number of cases per group, and g = number of groups.

Table 12. *Intraclass Correlation Coefficients (ICC) of Mathematics Measures Averaged Across Imputed Datasets*

	ICC			
	M	*SD*	Min	Max
TEMA	.08	.03	.04	.16
Counting Aloud	.04	.03	-.01	.12
Counting Objects	.03	.02	-.01	.11
Grouping	.10	.03	.02	.16
Measurement	.06	.03	-.02	.11
Number Naming	.07	.03	.02	.14
Pattern Recognition	.03	.03	-.03	.07

Separate hierarchical regression analyses were conducted for each criterion variable. Results were pooled across imputed datasets according to Rubin's (1987) rules (described in the Method section). The range of R^2 values across imputed datasets are reported in the tables below. Regression coefficients with *p*-values of .05 or less were identified as statistically significant. Pretest scores on the measure were entered in the first block. The following predictors were entered in the second block: age in months (grand mean centered), race (0 = *Caucasian*, 1 = *Other*), and gender (0 = *male*, 1 = *female*). Group (0 = *comparison*, 1 = Number*Fun*) was entered into the third block. Results of the regression analyses are presented below for each criterion.

TEMA-3. Results of the multiple regression predicting TEMA-3 post-test scores are presented in Table 13. TEMA-3 pretest scores were a significant predictor, accounting for 20% of the unique variance in scores (sr^2 = .10-.40), indicating that children who obtained higher scores on the pretest also received higher scores on the post-test. Of the demographic characteristics, age achieved significance and contributed most of the unique variance, 3% on average (sr^2 = .00-.12.). Condition was not a significant predictor of TEMA-3 post-test scores. After controlling for the other predictors, Condition contributed 2%, on average, of the variance in post-test scores (sr^2 = .00-.07; d = .33). As described in the Method,, effect sizes also were calculated on mean scores within and between groups. Large effects were observed between pretest and post-test scores for both the intervention (d = 1.07) and comparison (d = 2.02) conditions; however, a large effect (d = .98) favoring the comparison group was observed when comparing post-test scores between conditions.

Table 13. *Pooled Results of Multiple Regression Analyses for Test of Early Mathematics Ability – 3rd Edition*

Predictors	B	SE	t	p-value	R^2	ΔR^2
Step 1						
Pretest	.94	.14	6.89	<.01	.30 - .65	.30 - .65
Step 2						
Age	.34	.15	2.19	.03		
Race	-.87	2.38	-.37	.72		
Gender	1.32	1.62	.82	.42	.37 - .68	.01 - .14
Step 3						
Condition	-2.48	1.68	-1.48	.14	.38 - .70	.00 - .08

Counting Aloud. Pretest scores and age in months achieved significance when predicting post-test scores on the EARLI Counting Aloud probe (Table 14). As with the TEMA results, pre-test scores contributed the most unique variance, 9% on average (sr^2 = .01-.20), followed by age in months, which contributed 5% on average (sr^2 = .00-.14). Condition was not a significant predictor (sr^2 = .03 - .04; d = .05). Small effects were observed when pretest and post-test scores were examined for the intervention (d =.20) and comparison (d =.39) conditions; however, a nominal effect (d = .05) was found when comparing post-test scores between conditions.

Table 14. *Pooled Results of Multiple Regression Analyses for EARLI Counting Aloud*

Predictors	*B*	*SE*	*t*	*p*-value	R^2	ΔR^2
Step 1						
Pretest	.56	.19	2.96	<.01	.04 - .30	.04 - .30
Step 2						
Age	.04	.02	2.02	.04		
Race	-.21	.35	-.60	.55		
Gender	.11	.21	.52	.61	.12 - .34	.03 - .16
Step 3						
Condition	.03	.25	.12	.91	.12 - .34	.00 - .04

Counting Objects. Pretest scores and age achieved significance in predicting Counting Objects pretest scores (Table 15). On average, age in months contributed the most unique variance to the model (6%; sr^2 = .01-.16), followed by pre-test scores (5%; sr^2 = .00-.14). Condition was not a significant predictor (sr^2 = .00-.04; *d* = .29). Small effects were observed when pretest and post-test scores were examined for the intervention condition (*d* = . 42) and comparison (*d* = . 37) conditions; however, a small effect (*d* = .29) favoring the comparison group was found when comparing post-test scores between conditions.

Table 15. *Pooled Results of Multiple Regression Analyses for EARLI Counting Objects*

Predictors	*B*	*SE*	*t*	*p*-value	R^2	ΔR^2
Step 1						
Pretest	.24	.09	2.68	<.01	.03 - .26	.03 - .26
Step 2						
Age	.04	.02	2.31	.02		
Race	-.27	.32	-.85	.34		
Gender	.25	.23	1.09	.28	.11 - .35	.04 - .21
Step 3						
Condition	-.17	.23	-.75	.46	.12 - .36	.00 - .04

Grouping. Pre-test scores and age in months were found to be significant predictors of EARLI Grouping probe post-test scores (Table 16). Age in months made the largest contribution (11%) of the unique variance on average (sr^2 = .02-.23). Grouping pretest scores contributed 7% unique variance, on average (sr^2 = .00-.17) to pretest scores. Condition was not a significant predictor (sr^2 = .02-.03; *d* = .00). Medium effects were observed when pretest and post-test scores were examined for the intervention condition (*d* = .57) and comparison (*d* = .63) conditions; however, no effect (*d* = .00) was found when comparing post-test scores between conditions.

Table 16. *Pooled Results of Multiple Regression Analyses for EARLI Grouping*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.49	.14	3.43	<.01	.09 - .32	.09 - .32
Step 2						
Age	.06	.02	3.11	<.001		
Race	-.45	.33	-1.35	.18		
Gender	.33	.23	1.47	.14	.26 - .51	.05 - .27
Step 3						
Condition	.08	.25	.33	.74	.27 - .52	.00 - .03

Measurement. Pretest scores were the only significant predictor of Measurement post-test scores, contributing 6% unique variance, on average (sr^2 = .00-.14), to the prediction (Table 17). While not significant, race contributed 4% unique variance, on average (sr^2 = .00-.10), more than for any other outcome measure. Condition group was not a significant predictor (sr^2 = .00-.03; *d* = .14). small effects were observed when pretest and post-test scores were examined for the intervention condition (*d* = .23) and the comparison (*d* = .47) conditions; however, a nominal effect (*d* = .14) favoring the comparison group was found when comparing post-test scores between conditions.

Table 17. *Pooled Results of Multiple Regression Analyses for EARLI Measurement*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.44	.15	2.90	<.01	.02 - .25	.02 - .25
Step 2						
Age	.04	.02	1.89	.06		
Race	-.57	.36	-1.56	.12		
Gender	.20	.25	.79	.43	.08 - .40	.02 - .21
Step 3						
Condition	-.11	.24	-.45	.65	.08 - .42	.00 - .03

Number Naming. The only significant predictor of Number Naming post-test scores was Number Naming pre-test scores (Table 18). Pretest scores accounted for 22% unique variance in the equation, on average (sr^2 = .08-.38). Condition group was not a significant predictor (sr^2 = .01-.02; *d* = .07). Small effects were observed when pretest and post-test scores were examined for the intervention condition (*d* = .23) and comparison (*d* = .25) conditions; however, a nominal effect (*d* = .06) favoring the comparison group was found when comparing post-test scores between conditions.

Table 18. *Pooled Results of Multiple Regression Analyses for EARLI Number Naming*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.98	.21	4.77	<.01	.18 - .55	.18 - .55
Step 2						
Age	.02	.02	1.06	.29		
Race	.01	.25	.04	.97		
Gender	.04	.17	.23	.82	.24 - .55	.00 - .08
Step 3						
Condition	.01	.20	.07	.94	.24 - .56	.00 - .02

Pattern Recognition. Pretest scores had a significant relationship with Pattern Recognition post-test scores (Table 19) and contributed 7% unique variance on average (sr^2 = .00-.30). Condition was not a reliable predictor of Pattern Recognition post-test scores (sr^2 = .02-.03; *d* = .04). Nominal effects were observed when pretest and post-test scores were examined for the intervention condition (*d* = .06) and comparison (*d* = .13) conditions, and a nominal effect (*d* = .04) favoring the comparison group was found when comparing post-test scores between conditions.

Table 19. *Pooled Results of Multiple Regression Analyses for EARLI Pattern Recognition*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.59	.21	2.82	<.01	.02 - .43	.02 - .43
Step 2						
Age	.03	.02	1.34	.18		
Race	-.10	.34	-.31	.76		
Gender	.12	.25	.51	.61	.06 - .45	.00 - .03
Step 3						
Condition	.00	.26	.01	.99	.07 - .45	.00 - .03

Follow-up analyses. In order to determine the amount of variance in the models contributed by the Number*Fun* intervention without controlling for the other predictors, the entry of variables into the equation was reversed. Reversing the order of entry for the variables had little effect on the variance accounted for by condition variable on any measure (ΔR^2 range = .00-.08). Tables presenting the results of these analyses can be found in Appendices AC to AI.

Due to concerns regarding implementation fidelity, another series of regression analyses were run to explore if including classroom into the model improved the prediction of outcome measure scores. The classroom variable was effect coded so that each Number*Fun* classroom was a separate intervention condition, with comparison classrooms as the reference group. As with the previous models, pre-test scores were entered in the first block of the equation, followed by age in months (grand mean centered) in the second block. Because gender and race did not make meaningful or significant contributions in the original equations, those variables were excluded from the follow-up equations. The effects coded classroom variable was entered into the last block.

Prediction of post-test scores was improved by adding the effects coded classroom variable to the equation, while controlling for pre-test scores and age, for the TEMA-3 and the EARLI Number Naming probe (Tables 20 and 21, respectively), but not the other outcome measures (see Appendices AJ to AN). For the TEMA-3, half (50%) of the regression analyses of the imputed data sets indicated that adding the classroom variable significantly improved prediction at the $p < .05$ level. Being a child in Teacher 1's classroom was associated with a reduction of TEMA-3 post-test score of approximately 2.5 points (sr^2 = .00 - .04), whereas being in Teacher 2's classroom was associated with a gain of approximately two points (sr^2 = .00 - .04).

Table 20. *Pooled Results of Multiple Regression Analyses for Test of Early Mathematics Ability – 3rd Edition with Classroom Added as a Predictor*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.94	.14	6.89	<.01	.30 - .65	.30 - .65
Step 2						
Age	.32	.15	2.13	.03	.35 - .68	.00 - .12
Step 3						
Teacher 2	2.04	1.72	1.19	.24		
Teacher 3	-1.45	1.58	-.92	.36		
Teacher 1	-2.53	1.86	-1.36	.18	.39 - .72	.01 - .12

In 60.5% of imputed data sets, adding the classroom variable to the prediction of Number Naming post-test scores reached significance at the $p < .05$ level. A .40 reduction in scaled score points was associated with being in Teacher 1's classroom (sr^2 = .00 - .14). Being a child in Teacher 3's classroom was associated with a gain of .38 scaled score points (sr^2 = .00 - .11).

Table 21. *Pooled Results of Multiple Regression Analyses for EARLI Number Naming with Classroom Added as a Predictor*

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.98	.21	4.77	<.01	.18 - .55	.18 - .55
Step 2						
Age	.02	.02	1.07	.29	.25 - .55	.00 - .07
Step 3						
Teacher 2	-.04	.21	-.17	.87		
Teacher 3	.38	.21	1.81	.07		
Teacher 1	-.40	.29	-1.37	.18	.31 - .60	.01 - .15

Chapter 5

Discussion

The purpose of this study was to investigate the efficacy of a new early numeracy intervention, Number*Fun*. Three Head Start teachers implemented the intervention in their classrooms, while three other teachers engaged in usual practices (i.e., Creative Curriculum). Early mathematics measures were collected before and after the 13-week implementation period to determine if children who participated in Number*Fun* classrooms made greater gains in early mathematics skills than children in the comparison classrooms. Additional information was collected regarding the fidelity and acceptability of the intervention to provide a context for interpreting the outcomes obtained.

Summary of Key Findings Relative to Hypotheses

Child outcomes. Separate hierarchical regression analyses were run on the seven primary child outcome measures. Intervention condition was regressed on post-test scores after accounting for pre-test scores and demographic characteristics (i.e., age, race, and gender). Based on the results of these analyses, children in the intervention classrooms performed no differently than children in the comparison classrooms on measures of early mathematics skills. The only predictors that achieved significance were pretest scores, which predicted post-test scores on all of the outcome measures, and age in months, which predicted post-test scores of the TEMA-3, Counting Aloud, and Counting Objects, and Grouping, but not Number Naming, Pattern Recognition, or Measurement. Although not statistically significant, there was a small observed association between being in a

Number*Fun* classroom and having a slightly higher score on the Counting Aloud, Grouping, Pattern Recognition, and Number Naming EARLI Numeracy probes.

Adding gender and race into the equation did not significantly contribute to the prediction of post-test scores. While gender and/or ethnic group differences in mathematics achievement has been identified as early as kindergarten (e.g., Jordan et al., 2006; Lee, Austry, & Fox, 2008; Penner & Paret, 2008), few studies have examined these differences in children of preschool age. Arnold et al. (2002), found both gender and ethnic group differences in post-test scores on the TEMA-2. Ginsberg and Pappas (2004) and Clements et al. (2008), on the other hand, reported no gender or ethnic group differences in mathematics outcomes. The findings of the current study regarding gender differences are consistent with the majority of published findings.

Implementation fidelity. Interpreting the results obtained in the current study within the context of implementation fidelity may provide further insight into the efficacy of the Number*Fun* intervention.

Classroom observation indicated that teachers generally implemented the activities as they were presented in the manual and met activity objectives. Based on activity log entries, Teachers 1and 2 reported following the intervention schedule provided in the manual and were no more than a day or two off schedule at any given time. Teacher 3, however, frequently deviated from the sequence of activities provided in the manual. As a result, the children in her classroom completed only two-thirds of the Number*Fun* activities.

Teachers also reported facilitating the activities as they were described in the intervention manual with few modifications. Most modifications were minor and typically consisted of substituting materials or providing more support for some children in the classroom (e.g., providing hints/prompts, using hand-over-hand assistance with motor tasks). For example, one teacher substituted shaving cream for sand during an activity where children practiced writing numerals in the sand with their fingers. On a few occasions, Teacher 1 modified activities in ways that were inconsistent with activity objectives, thus changing the nature of the task. For example, instead of following the activity directions for the aforementioned number writing task, Teacher 1 hid small plastic bears in the sand, presented the children with a written numeral, and had them dig for that number of bears in the sand. On a few other tasks that required children to count pictures of objects on an index card and then find a match (either with the same number of pictures or with the corresponding written numeral), Teacher 1 modified the activity by having children count the pictures on the index card without finding a match.

Mathematics instruction in Number*Fun* versus comparison classrooms. Overall, Number*Fun* teachers reported spending nearly a third more instructional time on mathematics per day than comparison teachers. The majority of activities conducted in Number*Fun* and comparison classrooms targeted number and operations, followed by patterns/measurement and geometry/spatial sense. This skill focus is consistent with what other researchers have found when examining preschool teachers' mathematics instruction (Sarama et al., 2004). Number*Fun* teachers, however, targeted a wider array of number and operation skills relative to the comparison teachers, who primarily focused on counting and number recognition. Comparison teachers were more likely to conduct mathematics activities within the context of other activities (e.g., reviewing the calendar, identifying what happens first, second, third in a story, counting words written on a board, counting steps taken from classroom to restroom). Number*Fun* teachers, on the other hand, were more likely to conduct activities in which early mathematical concepts were the central focus. Number*Fun* Teachers also demonstrated more characteristics of effective and quality instruction than did comparison teachers.

There also were some similarities among Number*Fun* and comparison teachers. For instance, both groups of teachers were rated highly on measures of preparation and organization in facilitating mathematics activities during the classroom observations. In contrast, all teachers received lower ratings on behaviors associated with the promotion of mathematical thinking. Teachers across conditions were generally less successful at encouraging children to think and reflect upon the concepts being introduced or evaluate other children's answers, and they frequently failed to expand upon children's responses.

Number*Fun* acceptability. Overall, teachers reported having a positive experience with the intervention; however, there were indications that some aspects of the intervention were unsatisfactory. For example, all teachers expressed dissatisfaction with the length of activities. Upon follow-up, teachers stated that many transition activities were too long (an average of 8 minutes, but up to 15 minutes according to teacher logs). Teacher 1 also reported that many of the activities were not age-appropriate for the children in her classroom, and she indicated that more professional development would have been beneficial to implement the intervention. Examination of teacher ratings of specific activities revealed that Teachers 1 and 3 found some activities to be moderately difficult to implement. Consistent with her responses on the survey, Teacher 1 indicated that several activities were difficult for children in her classroom to complete. She also reported that the activities were less engaging than Teachers 2 and 3. If teachers did not implement the intervention as it was designed, examining teacher acceptability information may provide insight as to why this was the case.

Interpretation of Results Relative to Previous Research

Results of the regression analyses indicating that participation in the Number*Fun* intervention was not associated with greater gains in early mathematics skills than usual practice are surprising in light of findings from other similar studies. For example, Arnold et al. (2002) reported a large effect (d = 1.21) for their early mathematics intervention, despite the fact that the intervention was only 6 weeks in duration and teachers facilitated fewer activities than Number*Fun* teachers. Comparing the findings of Arnold et al.'s study to the current study is difficult due to the lack of information provided by the authors. No information was provided about the characteristics of the children within each condition and no information was provided about the curriculum, if any, used in the comparison classroom.

Other short-term inventions (3 - 8 weeks) also have demonstrated medium to large effect sizes (Pagani et al., 2006; Ramani et al., 2008; Young-Loveridge, 2004) but these interventions differed from Number*Fun* in four primary ways. First, each intervention targeted a small number of mathematics skills, whereas Number*Fun* was broad in scope. Focusing on a smaller skill set results in more practice and repetition for the children, which is important when learning a new skill. Second, two interventions (i.e., Ramani et al., Young-Loveridge) consisted of playing games (e.g., commercial or researcher-created board games, songs, books, manipulatives) with individual or small groups of children (2 to 3) working with an adult (e.g., teacher or researcher), whereas most of Number*Fun* activities occurred in a large or whole group setting. Working with one or more child at a time makes it easier for a teacher to monitor the child's learning and provide frequent feedback, things that are more difficult in a large group setting. Third, three of the four short-term studies (i.e., all but Arnold et al.) used assessments developed specifically for the study, thus making it more likely to find positive gains. This was true for long-term interventions as well (Clements et al., 2008; Sophian, 2004; Starkey et al., 2004). Finally, two of the studies used gain scores in their analyses, which might have led to overestimated effect sizes (Matton et al., 2009).

Working Hypotheses for Outcomes

Several factors may contribute to the lack of evidence supporting efficacy of the Number*Fun* intervention: measurement error, implementation fidelity, and Number*Fun* activities. These factors are discussed in more detail below.

Measurement error. Measurement error in this study may come from two potential sources: the measures used to assess early mathematics skills and the characteristics of the children. The TEMA-3 is the only published comprehensive mathematics assessment that can be completed by children as young as 36 months; however, the TEMA-3 has several limitations relevant to the current study. First, the normative sample underrepresented children from families living in poverty, including less than one-hundred 3- and 4-year-olds. Second, little information is presented in the examiner’s manual regarding the psychometric properties of the measure with 3- and 4-year-olds. Third, there appears to be floor effects regarding the difficulty of the items. In the current study, less than half (48%) of the sample (before imputation) answered five or more questions correctly at the pre-test (range: 0 – 25) and less than 50% answered 10 or more questions correctly on the post-test (range: 0 – 42).

The EARLI Numeracy probes also are not without limitations. While developed specifically to assess numeracy skills in 3- and 4-year-olds enrolled in Head Start classrooms, the probes norms have yet to be established. Additional data needs to be collected on a larger, more representative sample of preschoolers. In addition, the latest analyses (DiPerna et al., 2009) suggest that the probes are likely most appropriate for screening purposes.

Characteristics of preschool-aged children (e.g., language proficiency, distractibility, eagerness to please examiner) and situational variables (e.g., unfamiliar adult, lack of prior testing experience) are additional factors that often make it difficult to obtain reliable estimates of young children's skills using traditional assessment methods (Nagle, 2007). These factors may lead to unstable test scores and potentially introduce another source of variation not accounted for in the regression model.

Of particular concern is the age distribution of the children across intervention conditions. Intervention classrooms had significantly more 3-year-old children (i.e., those not age-eligible for kindergarten the subsequent school year) than the comparison classrooms. Three-year-olds, as a group, scored significantly lower than 4-year-olds on all measures. In addition, 3-year-olds made significant gains on two of the seven measures; whereas 4-year-olds made significant gains on four measures (and gains on two other measures approached but did not meet the significance criterion of .05). This also was reflected in the increased variance in test scores from pre-test to post-test. It appears that either 3-year-olds, in particular, did not benefit from the intervention due to their cognitive immaturity, or measuring the developing cognitive skills of 3-year-olds is especially problematic and subject to error. In fact, of the early mathematics intervention studies discussed in the literature review, only two (Sophian, 2004; Arnold et al. 2002) included 3-year-olds in their study, with mixed results. Significant age differences were found in Sophian's sample, but not Arnold et al.'s.

Implementation fidelity. Classroom observation data suggested that teachers generally implemented the intervention as designed. However, examination of the activity logs revealed that they might not have been completed accurately. First, logs were to be completed on a daily basis; however, examination of Teacher 3's log at the end of a classroom observation revealed that her last entry was recorded 3 weeks prior. Second, the log entry of Teacher 1 on Day 39 of the intervention did not match the records from the classroom observation. That is, Teacher 1 reported conducting the three Number*Fun* activities scheduled for that day; however, she was observed facilitating three non-Number*Fun* activities during the classroom visit, which she did not record in her log. Teacher 1's log entries for the first several weeks of the study also were of concern. Based on how she numbered the activities and recorded her entries, it appeared that she might have only conducted one of the three daily activities but entered information about the other activities later. Finally, during a meeting with teachers at the conclusion of the study, all indicated that there were several times when they could not facilitate all three activities in one day, but this was not reflected in their logs. If the logs are unreliable, determining the efficacy of the Number*Fun* intervention is difficult due to the potential variation in implementation among teachers.

There also were observable differences between the Number*Fun* teachers in terms of quality of instruction. For instance, Teacher 1 was observed to make fewer attempts to engage children in instruction. She also often failed to provide multiple opportunities for children to practice target skills, to check for understanding, and to correct misunderstanding. These types of instructional behaviors were emphasized in Number*Fun* activities.

Based on the teacher acceptability data, it appears as though some aspects of the activities may explain variation in implementation fidelity. Specifically, Teacher 1 rated activities as slightly more difficult for children, on average, than other teachers. Additionally, her logs indicated several substantial modifications to the activities. She also indicated that more professional development was needed. It may be that she lacked an understanding of the objectives of the activities, which first, made it more difficult for her to effectively facilitate the activities, and second, made her more likely to modify the activities in ways that were inconsistent with the original objectives. In addition, all of the teachers indicated that the activities were too long, especially transition activities. This may have led them to either facilitate fewer transition activities, or rush through them.

Due to the concerns raised regarding implementation fidelity and instructional quality, additional analyses were run with pre-test scores, age in months, and intervention condition regressed on post-test scores. These analyses yielded interesting results, suggesting that classroom might be a significant predictor of TEMA-3 and Number Naming post-test scores. Being a child in Teacher 1's classroom was associated with a decrease of nearly 2.5 points on the TEMA-3 and .40 scaled score points on Number Naming controlling for the other predictors. Being a child in Teacher 2's classroom was associated with a gain of 2 points on the TEMA-3, whereas being a child in Teacher 3's classroom was associated with a decrease of nearly 1.5 points on the TEMA-3 but a gain of .40 scaled score points on the Number Naming probe.

Intervention activities. While measurement error and teacher implementation fidelity may have weakened the effects of the intervention, it may be that the intervention was simply not effective at increasing children's performance on mathematics measures beyond that produced by usual practice. Three aspects of the intervention may be responsible for the outcomes: timing, duration, and pace. First, the intervention occurred during the second half of the academic year. It may be that the activities facilitated early on in the intervention targeted skills already covered in the classroom. This is partially supported by teacher reports. Teachers were more likely to rate activities that occurred earlier in the intervention as “too easy” and more likely to utilize upward extension modifications.

Second, the intervention spanned a 13-week period during the latter part of the academic year. Although other published studies of shorter duration have demonstrated positive effects (e.g., Ramani et al., 2008; Siegler et al., 2009; Young-Loveridge, 2004), these interventions were narrowly focused, whereas Number*Fun* targets a wider array of early mathematics skills. It may be that more time is required to detect the effects of an intervention with a broader focus.

Finally, the pace at which new activities were introduced may have been too quick. Most activities were repeated once. Math center activities were available for independent exploration for 5 days. Children, however, may have benefited more from additional practice, particularly with the more advanced concepts. The need for more time spent on activities is partially supported by teacher report. Teacher 3 often repeated activities more times than directed by the intervention manual, which resulted in her omitting one-third of the activities.

Limitations

Several limitations need to be considered when interpreting the results of the current study. Due to limited resources, a small number of volunteer classroom teachers from one Head Start program were recruited to participate in this study. The number of participating classrooms, the age group of the children taught in the classrooms, and the location of the classrooms prevented the researcher from randomly assigning classrooms to each condition. Instead, classrooms were matched based on the age group of the children in the classroom and then classrooms were randomly assigned to conditions. The sampling procedures limit the generalizability of the results to classrooms other than the ones in this study.

Additionally, the small sample size resulted in a substantial loss of power. Power to detect a meaningful R^2 change in each analysis ranged from .08 to .29. Given the size of this sample, it is estimated that a nearly medium effect ($f^2 = .14$) could have been detected ($\alpha = .05$, power = .80). This suggests that if the small effects ($sr^2 = .01$ to .04) found in this study are similar to population levels, then the significance tests were not sensitive enough to detect differential impact of intervention status on outcome measures due to a lack of power.

Another limitation of this study is that the analyses conducted did not take into consideration the fact that children were nested within classrooms. Children in the same classroom may share similar experiences that, in turn, may cause their test scores to be more correlated with those of classmates than those of children in other classrooms. A method such as Hierarchical Linear Modeling (HLM), which requires a larger sample and is more powerful (Raudenbush & Bryk, 2002), would have been more appropriate to answer the questions posed in this study. While the estimates produced by the analyses that ignore clustering are not biased, standard errors (and subsequently significance tests that rely on standard errors) may be biased. Although the intraclass correlations reported in the current study were relatively small, the results must be interpreted with caution.

An additional limitation of this study is the amount of missing data due to attrition and child absence. Multiple imputation has been identified as best practice for estimating missing data values. While multiple imputation is a substantial improvement over traditional missing data techniques in producing unbiased estimates, it is not a substitute for complete data and is limited by the properties of the original data (Graham & Schafer, 1999).

Given issues regarding implementation fidelity, it is apparent that classroom observations needed to occur more frequently and for longer periods. Limited resources and practical constraints restricted classroom visits to once a month, increasing reliance on teacher completed activity logs. Suspected inaccuracies in teacher reports, however, prevent firm conclusions regarding the degree to which the intervention was implemented with fidelity. Besides providing additional fidelity information, more classroom visits would have provided the researcher with additional opportunities to model instructional strategies and provide coaching as necessary.

Efforts were made to ensure that intervention and comparison classrooms were not located in the same buildings; however, after the teachers had been assigned to their classrooms, Head Start administrators relocated two classrooms. This caused two intervention and one comparison classrooms to be in the same building. Additionally, when the replacement teacher of Classroom 6 (which originally had been assigned to the intervention condition) declined to participate in the study the classroom her classroom was reassigned to the comparison condition. This then lead her classroom to be in the same building as an intervention classroom. Even though intervention teachers were instructed not to discuss the intervention with comparison teachers, it appears that some discussion may have occurred. For instance, materials for one Number*Fun* activity were found in a comparison classroom during a classroom observation. In addition, it was observed on separate occasions that intervention teachers displayed products of Number*Fun* activities on the walls inside and outside of their classrooms, in full view of comparison teachers.

Future Directions for Research

Although the results obtained were not consistent with predictions, additional research on the Number*Fun* intervention is warranted as evidence of its effectiveness remains to be established. Specifically, it is unclear if the lack of improved outcomes in intervention classrooms compared to comparison classrooms is due to: (a) the intervention was no more effective than business as usual, (b) the intervention was not implemented as designed, or (c) the intervention was not appropriate for 3-year-olds. As such, future research should first employ more intensive fidelity procedures. One step would be to conduct implementation fidelity observations more frequently, perhaps on a weekly basis. Additionally, a more specific fidelity checklist could be developed to assess teacher's adherence to the manual while facilitating intervention activities. Second, a follow-up study should be conducted to examine the impact of Number*Fun* exclusively on 4-year-old children.

Once fidelity of implementation issues are corrected, further research efforts should consider modifying aspects of the intervention. Specifically, the duration of the intervention should be lengthened to span the academic year, which would allow teachers to target specific skills with greater depth and provide children with more opportunities to practice the skills they are acquiring. In addition, teachers reported that transition activities were too long. These types of activities need to be examined further to determine if they can be shortened without compromising the objectives of the activities or if teachers need more guidance on how to incorporate mathematics activities successfully into transition times.

Teacher reports indicated the need for more comprehensive and ongoing professional development program to accompany the intervention. Based on the results from the present study, teachers may have benefitted from professional development that placed a greater emphasis on increasing their understanding of early numeracy skills and supporting their ability to provide effective mathematics instruction. In addition, teachers may have benefitted from regular meetings with the other intervention teachers to discuss issues regarding implementation, questions about content, or mathematics instruction related topics selected by the researcher. Future studies could investigate whether or not providing comprehensive professional development in conjunction with the intervention has an indirect effect on child outcome measures.

Although teachers reported that implementation of the intervention was feasible, they were provided with activity materials in advance and nominal advanced preparation of materials of materials was required to facilitate the activities. In future studies, particularly with larger sample sizes, preparing teacher's materials in advance may not be feasible for the researcher. Future research should assess if teachers continue to find the intervention feasible if they are responsible for preparing their own activity materials.

Conclusion

Children living in poverty are at greater risk for experiencing academic learning difficulties than their peers living in families with greater economic means. Discrepancies in mathematics skills are evident before children enter formal schooling, so it is important to intervene early to decrease the likelihood of future mathematics difficulties. Programs such as Head Start aim to foster school readiness in preschool-aged children; however, recent national evaluations of the program have indicated it is having no discernible effect on children's mathematics skills. In light of these findings, Head Start programs need access to high quality curricula that promote early mathematics development in young children.

The purpose of the current study was to assess the efficacy of a newly developed early mathematics intervention, Number*Fun*, in Head Start classrooms. Children who participated in Number*Fun* made significant gains in early mathematics skills; however, these gains were not beyond those of children in comparison classrooms where teachers utilized the regular Head Start curriculum (Creative Curriculum). Although, intervention teachers reported implementing the intervention as it was designed; however, inconsistencies in reporting were evident in two teachers' logs. Examination of classroom observations and teacher activity logs indicated that Number*Fun* teachers spent more time on mathematics in their classroom and generally received higher quality of instruction ratings; but, the quality of mathematics instruction was variable across intervention teachers.

Concerns regarding implementation fidelity, as well as methodological limitations of the study, hinder drawing firm conclusions about the efficacy of the Number*Fun* intervention. Future research needs to improve upon the procedures used to measure intervention fidelity and focus on developing a comprehensive professional development program to accompany Number*Fun* in order to support teachers in fostering early numeracy skills in preschoolers. Finally, the utility of conducting Number*Fun* activities during transition periods needs to be reexamined to determine if those activities should be shortened or facilitated at a different time of day.

REFERENCES

Adelman, C. (1999). *Answers in the toolbox: Academic intensity, attendance patterns and bachelor's degree attainment.* Washington, DC: U. S. Department of Education.

Administration for Children and Families. (2000). *Head Start child outcomes framework*. Retrieved May 15, 2006 from http://www.hsnrc.org/CDI/pdfs/UGCOF.pdf

Administration for Children and Families. (2002). *Head Start history*. Retrieved December 16, 2005, from http://acf.hhs.gov/programs/hsb/about/history.htm

Allison, P. D. (2002). *Missing data*. (Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-136). Thousand Oaks, CA: Sage.

Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math development in Head Start classrooms. *Journal of Educational Psychology, 94*, 762-770.

Aunola, K., Leskinen, E., Lerkkanen, M. K., & Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology, 96*, 699-713.

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type 1 errors? Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology, 73*, 924-935.

Balfanz, R. (1999). Why do we teach children so little mathematics? Some historical considerations. In J. V. Copley (Ed.), *Mathematics in the early years* (pp. 3-10). Reston, VA: National Council of Teachers of Mathematics.

Baroody, A. J., Lai, M., & Mix, K. S. (2006). The development of young children's early number and operation sense and its implications for early childhood education. In B.

Spodek & O. N. Saracho (Eds.), *Handbook of research on the education of young children* (10th ed., pp 187-221). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Berch, D. B. (2005). Making sense of number senses: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*, 333-339.

Blevins-Knabe, B., & Musun-Miller, L. (1996). Number use at home by children and their parents and its relationship to early mathematical performance. *Early Development and Parenting, 5*, 35-45.

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling, 15*, 651-675.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*, 223-240.

Case, R. (1998, April). *A psychological model of number sense and its development*. Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Case, R., Griffin, S., & Kelly, W. M. (1999). Socioeconomic gradients in mathematical ability and their responsiveness to intervention during early childhood. In D.P. Keating & C. Hertzman (Eds.). *Developmental health and wealth of nations*. New York: Guilford Press.

Clements, D. H. (2004). Major themes and recommendations. In D. Clements, J. Sarama, and A. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 7-72). Mahwah, NJ: Erlbaum.

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*, 443-494.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cooper, R. G., Jr. (1984). Early number development: Discovering number space with addition and subtraction. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 157-192). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* New York, NY: Oxford University Press.

Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade*. Washington, DC: National Center for Education Statistics.

Developing Skills Checklist. (1990). Monterey, CA: CBT/McGraw-Hill.

Dougherty, C. (2003). Numeracy, literacy and earnings: Evidence from the National Longitudinal Survey of Youth. *Economics of Education Review, 22,* 511-521.

DiPerna, J. C., & Morgan, P. (2006). EARLI Numeracy Probes. University Park, PA: The Pennsylvania State University.

DiPerna, J. C., Morgan, P., Lei, P. W. (2009). Early Arithmetic and Reading Learning Indicators project: Final report. University Park, PA: The Pennsylvania State University.

DiPerna, J. C., Morgan, P. L., Lei, P. W., Reid, E. E., & Wu, Q. (2006, November). *The EARLI Mathematics probes: Initial reliability and validity evidence for children in Head Start.* Poster session presented at annual meeting of the National Association for the Education of Young Children, Washington, DC.

Dorman, J. P. (2008). The effect of clustering on statistical tests: an illustration using classroom environment data. *Educational Psychology, 28,* 583-595.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446.

Elkind, D. (1981). *The hurried child*. Redding, MA: Addison Wesley.

Evan, A., Gray, T., & Olchefske, J. (2006). *The gateway to student success in mathematics and science.* Washington, DC: American Institutes for Research.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for mathematical disabilities. *Journal of Experimental Child Psychology, 74*, 213-239.

Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 33*, 18-28.

Ginsburg, H. P., & Baroody, A. J. (2000). Test of Early Mathematics Ability-Second Edition. Austin, TX: Pro-Ed.

Ginsburg, H. P., & Baroody, A. J. (2003). Test of Early Mathematics Ability-Third Edition. Austin, TX: Pro-Ed.

Ginsburg, H. P., & Pappas, S. (2004). SES, ethnic, and gender differences in young children's informal addition and subtraction: A clinical interview investigation. *Applied Developmental Psychology, 25*, 171-192.

Ginsburg, H. P., Cannon, J., Eisenband, J., & Pappas, S. (2006). Mathematical thinking and learning. In K. McCartney & D. Phillips (Eds.), *Blackwell Handbook of Early Childhood Development* (pp. 208-229). Malden, MA: Blackwell.]

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context* (NCES 2009–001Revised). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576.

Graham, T. A., Nash, C., & Paul, K. (1997). Young children's exposure to mathematics: The childcare context. *Early Childhood Education Journal, 25*, 31-38.

Graham, T. A., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206-213.

Graham, J. W., & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 1-29). Thousand Oaks, CA: Sage.

Griffin, S. (2003). The development of math competence in the preschool and early school years: Cognitive foundations and instructional strategies. In J. Royer (Ed.), *Mathematical cognition.* Greenwich, CT: Information Age Publishing.

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly, 19,* 173-180.

Grigg, W., Donahue, P., & Dion, G. (2007). *The nation's report card: 12th-grade reading and mathematics 2005* (NCES 2007-468). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U. S. Government Printing Office.

Halberda, J., Mazzocco, M. M. M., Feigenson, L. (September 7, 2008). *Individual differences in non-verbal number acuity correlate with maths achievement.* Retrieved from www.nature.com.

Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics, 32,* 151-179.

Hoyle, R. H., Georgesen, J. C., & Webster, J. M. (2001). Analyzing data from individuals in groups: The past, the present, and the future. *Group Dynamics: Theory, Research, and Practice, 5,* 41-47.

Jeličić, H., Phelps, E., Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: the persistence of bad practices in developmental psychology. *Developmental Psychology, 45,* 1195-1199.

Jordan, N. C., Kaplan, D., Oláh, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77,* 153-175.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, *99*, 422-431.

Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48,* 339-348.

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology, 83,* 126-137.

Kilbanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 4*, 59-69.

Klein, A., & Starkey, P. (2004). Fostering preschool children's mathematical knowledge: Findings from the Berkeley Math Readiness Project. In D.H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 343-360). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Lee, J., Austry, M. M., & Fox, J. (2008). Investigating children's mathematics readiness. *Journal of Research in Childhood Education, 22*, 316-328.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

LeFevre, J. A., Clarke, T., & Stringer, A. P. (2002). Influences of language and parental involvement on the development of counting skills: Comparisons of French- and

English-speaking Canadian children. *Early Child Development and Care, 172*, 283-300.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lipton, J., & Spelke, E. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science, 14,* 396-401.

Malofeeva, E. V., Ciancio, D., & Day, J. D. (2001, August). *Development of emergent math and literacy skills.* Paper presented at the annual meeting of the American Psychological Association, San Francisco, CA.

Matton, N., Vautier, S., & Raufaste, E. (2009). Situational effects may account for gain scores in cognitive ability testing: A longitudinal SEM approach. *Intelligence, 37,* 412-421.

Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. Learning Disabilities Research and Practice, 20, 142-155.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30-46.

Mitra, A. (2002). Mathematics skill and male-female wages. *Journal of Socio-Economics, 31*, 443-456.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2007). *TIMSS 2007 assessment frameworks*. Boston, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management, 19*, 547-568.

Murnane, R. J., Willett, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics, 77*, 251-266.

Nagel, R. J. (2007). Issues in preschool assessment. In B. Bracken & R. Nagel (Eds.), *Psychoeducational assessment of preschool children* (4th ed., pp. 29-48). Mahwah, NJ: Erlbaum.

National Center for Education Statistics. (2009). *The nation's report card: Mathematics 2009* (NCES 2010-451). Institute of Education Sciences, U. S. Department of Education, Washington, D. C.

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

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8: A quest for coherence*. Retrieved from http://www.nctmmedia.org/cfp/full_document.pdf

National Education Goals Panel. (1999). *The National Education Goals report: Building a nation of learners, 1999*. Washington, DC: U.S. Government Printing Office.

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

National Mathematics Advisory Panel. (2008b). *Foundations for success: Reports of the task groups and subcommittees*. Washington, D. C.: U. S. Department of Education.

National Science Board. (2008). *Science and engineering workforce: Realizing America's potential.* Arlington, VA: National Science Foundation (Vol. 1, NSB 08-01; Vol. 2, NSB 08-01A).

Okamoto, Y., & Case, R. (1996). Exploring the microstructure in the domain of number. In R. Case & Y. Okamoto (Eds.), the role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development, 60*, 27-58.

Pagani, L. S., Jalbert, J. & Girad, A. (2006). Does preschool enrichment of precursors to arithmetic influence intuitive knowledge of number in low income children? *Early Childhood Education Journal, 34*, 133-146.

Pasnak, R., Greene, M. S., Ferguson, E. O., & Levit, K. (2006). Applying principles of development to help at-risk preschoolers develop numeracy. *The Journal of Psychology, 140*, 155-173.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction.* (3rd ed.). Fort Worth, TX: Harcourt Brace.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*, 239-253.

Pennsylvania Department of Education and Public Welfare. (2007). *Pennsylvania learning standards for early childhood*. Harrisburg, PA: Pennsylvania Department of Education.

Peters, S. (1998). Playing games and learning mathematics: The results of two intervention studies. *International Journal of Early Years Education, 6*, 49-58.

Piaget, J. (1965). *The child's conception of number*. New York, NY: Norton.

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Reid, E. E., Morgan, P. L., DiPerna, J. C., & Lei, P. W. (2006). Development of measures to assess young children's early academic skills: Preliminary findings from a Head Start-University partnership. *Insights on Learning Disabilities, 3*, 25-38.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Saracho, O. N., & Spodek, B. (2008). Trends in early childhood mathematics research. In O. Saracho & B. Spodek (Eds.*), Contemporary perspectives on mathematics in early childhood education* (pp. vii-xx). Charlotte, NC: Information Age Publishing, Inc.

Sarama, J., DiBiase, A., Clements, D. H., & Spitler, M. E. (2004). The professional development challenge in preschool mathematics. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 415-446). Mahwah, NJ: Lawrence Erlbaum Associates.

Saxe, G. B., Guberman, S. R., Gearhart, M. (1987). Social processes in early number development. *Monographs of the Society for Research in Child Development, 52*.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33*, 545-571.

Seo, K. H., & Ginsburg, H. P. (2004). What is developmentally appropriate in early childhood mathematics education? Lessons from new research. In D. H. Clements &

J. Sarama (Eds.). *Engaging young children in mathematics* (pp. 91-104). Mahwah, NJ: Erlbaum.

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

Siemer, M., & Joorman, J. (2003). Power and measures of effect size in analysis of variance with fixed versus random nested factors. *Psychological Methods, 8,* 497-517.

Skwarchuk, S. (2009). How do parents support preschoolers' numeracy learning experiences at home? *Early Childhood Education Journal, 37,* 189-197.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Sophian, C. (2004). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly, 19,* 59-81.

Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43,* 93-126.

Starkey, P., Klein, A., Chang, I., Dong, Q., Pang, L., & Zhou, Y. (1999, April). *Environmental supports for young children's mathematical development in China and the United States.* Paper presented at the meeting of the Society for Research in Child Development, Albuquerque, NM.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19,* 99-120.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Creative Curriculum for Preschool®. (2010). Bethesda, MD: Teaching Strategies, Inc. Retrieved from http://www.teachingstrategies.com/page/CCPS_Overview.cfm.

U.S. Department of Health and Human Services, Administration for Children and Families. (June 2005). *Head Start Impact Study: First year findings*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Children and Families. (February 2006a). *Family and Child Experiences Survey (FACES 2000): Technical report*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Children and Families. (December 2006b). *FACES 2003 research brief: Children's outcomes and program quality in Head Start*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Children and Families. (June 2005). *Head Start Impact Study: First year findings*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Children and Families. (June 2010). *Head Start Impact Study: Final Report*. Washington, DC.

Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: Evidence from manual search. *Journal of Cognition and Development, 1,* 249-280.

Wampold, B. E., & Serlin, R. C. (2000). The consequences of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods, 5,* 425-433.

Williamson, G. L., Appelbaum, M., & Epanchin, A. (1991). Longitudinal analyses of academic achievement. *Journal of Educational Measurement, 28,* 61-76.

Wolf, F.M. (1986). *Meta-analysis: Quantitative methods for research synthesis.* London: Sage.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Young-Loveridge, J. M. (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly, 19*, 82-98.

Appendix A

Head Start Child Outcomes Framework – Mathematics (2000)

Domain/Indicator

- 3.1: Number & Operations
  - 3.1.1: Demonstrates increasing interest and awareness of numbers and counting as a mean for solving problems and determining quantity.
  - 3.1.2: Begins to associate number concepts, vocabulary, quantities, and written numerals in meaningful ways.
  - 3.1.3: Develops increasing ability to count in sequence to 10 and beyond.
  - 3.1.4: Begins to make use of one-to-one correspondence in counting objects and matching groups of objects.
  - 3.1.5: Begins to use language to compare numbers of objects with terms such as more, less, greater than, fewer, equal to.
  - 3.1.6: Develops increased abilities to combine, separate, and name "how many" concrete objects.
- 3.2: Geometry & Spatial Sense
  - 3.2.1: Begins to recognize, describe, compare, and name common shapes, their parts, and attributes.
  - 3.2.2: Progresses in ability to put together and take apart shapes.
  - 3.2.3: Begins to be able to determine whether or not two shapes are the same size and shape.
  - 3.2.4: Shows growth in matching, sorting, putting in a series, and regrouping objects according to one or two attributes such as color, shape, or size.
  - 3.2.5: Builds an increasing understanding of directionality, order, and positions of objects, and words such as up, down, over, under, top bottom, inside, outside, in front, and behind.
- 3.3: Patterns & Measurement
  - 3.3.1: Enhances abilities to recognize, duplicate, and extend simple patterns using a variety of materials.
  - 3.3.2: Shows increasing abilities to match, sort, put in a series, and regroup objects according to one or two attributes such as shape or size.
  - 3.3.3: Begins to make comparisons between several objects based on a single attribute.
  - 3.3.4: Shows progress in using standard and non-standard measures for length and area of objects.

Appendix B

National Council of Teachers of Mathematics Curriculum Focal Points and Connections/Expectations (2006)

- Curriculum Focal Points/Expectations
  - Number and Operations: Developing an understanding of whole numbers, including concepts of correspondence, counting, cardinality, and comparison.
    - NCTM 1.1: Count with understanding and recognize “how many” in sets of objects.
  - Geometry: Identifying shapes and describing spatial relationships.
    - NCTM 2.1: Recognize, name, build, draw, compare, and sort two-and three-dimensional shapes.
    - NCTM 2.2: Describe attributes and parts of two- and three-dimensional shapes.
    - NCTM 2.3: Describe, name, and interpret relative positions in space and apply ideas about relative position.
    - NCTM 2.4: Describe, name, and interpret direction and distance in navigating space and apply ideas about direction and distance
    - NCTM 2.5: Find and name locations with simple relationships such as “near to” and coordinate systems such as maps.
    - NCTM 2.6: Create mental images of geometric shapes using spatial memory and special visualization
    - NCTM 2.7: Recognize geometric shapes and structures in the environment and specify their location.
  - Measurement: Identifying measurable attributes and comparing objects by using these attributes.
    - NCTM 3.1: Recognize the attributes of length, volume, weight, area, and time.
    - NCTM 3.2: Compare and order objects according to these attributes.
- Connections to the Pre-K Focal Points/Expectations
  - Data Analysis
    - NCTM 4.1: Sort and classify objects according to their attributes and organize data about the objects.
  - Algebra
    - NCTM 5.1: Sort, classify, and order objects by size, number, and other properties.
    - NCTM 5.2: Recognize, describe, and extend patterns such as sequences of sounds and shapes or simple numeric patterns and translate from one representation to another.
    - NCTM 5.3: Describe qualitative change, such as a student’s growing taller.

Appendix C

Pennsylvania Department of Education Logical Mathematics Early Learning Standards (2007)

Standard/Indicator

- Numbers: Learn about numbers, numerical representation, and simple numerical operations
    - LM 1.1: Use counting and numbers as part of play and as a means for determining quantity.
    - LM 1.2: Count up to ten objects in meaningful context with emerging one-to-one correspondence.
    - LM 1.3: Understand number concepts, vocabulary, quantities, and written numerals in meaningful ways.
    - LM 1.4: Begin to use language to compare numbers of objects with terms such as more, less, greater than, fewer, equal to.
    - LM 1.5: Develop increased abilities to combine, separate, and name "how many" concrete objects.
    - LM 1.6: Use ordinal number words to describe the position of objects (first, second, third).
    - LM 1.7: Begin to solve problems using numbers.
- Patterns: Understand patterns, relations, and functions
    - LM 2.1: Recognize how things are alike (comparisons) and identify objects that belong together (classification).
    - LM 2.2: Sort, categorize, classify, and order objects by one attribute.
    - LM 2.3: Sort, categorize, classify, and order objects by more than one attribute.
    - LM 2.4: Order objects by properties (e.g., from small to large, lightest to darkest).
    - LM 2.5: Explain why and how objects are organized.
    - LM 2.6: Recognize, describe, and extend patterns.
- Space/Shape: Develop concepts of space and shape
    - LM 3.1: Recognize, name, describe, build, draw, and compare two- and three-dimensional shapes.
    - LM 3.2: Put together and take apart increasingly more difficult shapes.
    - LM 3.3: Determine whether or not two shapes are the same size and shape.
    - LM 3.4: Recognize geometric shapes in books, artwork, and the environment.
    - LM 3.5: Understanding directionality, order, and positions of objects, using words such as up, down, over, under, top, bottom, inside, outside, in front of, behind.
    - LM 3.6: Shows an awareness of symmetry.
- Measurement: Develop and use measurement concepts
    - LM 4.1: Demonstrate awareness of measurement attributes (length, volume, weight, area, time, and temperature).
    - LM 4.2: Develop an awareness of seriation. Compare attributes such as length (shorter – taller), size (bigger – smaller), weight (heavier – lighter), in everyday situations.
    - LM 4.3: Use standard and non-standard measures in everyday situations.

Appendix C (continued)

- Data: Represent and interpret data
  - LM 5.1: Gather information about themselves and their surroundings.
  - LM 5.2: Contribute data for simple graphs.
  - LM 5.3: Organize and display data on graphs using objects and pictures.
  - LM 5.4: "Read" and interpret displays of data using words to compare (e.g., quantity, size, speed, and weight).
- Problem Solving: Reason, predict, and problem solve
  - LM 6.1: Make predictions based on observations and information.
  - LM 6.2: Use simple strategies to problem solve.
  - LM 6.3: Tell others how to solve a problem.
  - LM 6.4: Understand that there is more than one way to solve a problem
  - LM 6.5: Develop the ability to compare/contrast solution strategies
  - LM 6.6: Use the language of mathematics to express mathematical ideas.

Appendix D

Teacher Recruitment Letter

Number*Fun*: Promoting Early Numeracy Skill Growth in Head Start Children (The Number*Fun* Project)

1. **What it is:** Number*Fun* is a government funded research study. The primary goal of the research project is to develop practical and effective early numeracy instructional activities for use in Head Start Classrooms. The research study is being conducted by researchers (Erin Reid and James DiPerna) at Penn State and has been funded for two years (2007 - 2009) by the U. S. Department of Health and Human Services.

2. **What the intervention will look like**: Teachers who implement the intervention will lead classroom activities that have been shown to improve preschool-age children's early numeracy skills. The intervention activities are developmentally sequenced and designed to foster early numeracy skills through structured experiences and explicit instruction. The children will be engaged in two or more early numeracy activities (20-30 minutes) per day.

3. **How the effectiveness of the intervention will be evaluated:** The Number*Fun* intervention will be evaluated in two ways. First, research staff will test the early numeracy skills of participating children in the classroom on a regular basis using the *Test of Early Mathematics Ability, 3rd Edition,* and the *EARLI Numeracy Probes*. Both of these measure children's understanding of numbers, counting, and grouping. Second, you will be asked to keep a record of the early numeracy activities that occur in the classroom. In addition, research staff will come to your classroom once a month and observe these activities.

4. **What's involved for children in your class:** Measures will be collected five times a year. During the 2007-2008 school year, data will be collected twice in January and twice in May. At these times, Penn State research staff will work for approximately 20 minutes with each participating child in your classroom. During this brief session, the research assistant will ask the child to complete measures of their knowledge of letters, numbers, sounds, and sequence.

5. **What's asked of parents:** Letters will be mailed to parents informing them about the project and asking if they agree to have their child participate. Parents are asked to return the signed forms to the investigators only if they do *not* want their child to participate in the study.

6. **What's asked of you:** If you choose to participate, you will be asked to sign a consent letter and complete a brief questionnaire about your educational background, years of experience, etc. You will maintain a brief daily log of the numeracy related activities you lead in your classroom. Once a month, trained research staff will observe in your classroom. You also will be asked to complete surveys regarding the early academic skills and behaviors of the participating children in your classroom. In addition, some teachers will be selected to implement the program beginning in the spring of 2008. The remaining teachers will receive training on the program at the end of the project (late spring 2009) If you are selected as a teacher who will implement the intervention in her classroom this spring, you will have the

Appendix D (continued)

opportunity to share your ideas and provided feedback to the investigators as the intervention is being developed. You will also attend training sessions and maintain contact with the investigators on a bi-weekly basis. All participating teachers will be compensated $25 an hour for participating in the activities described in this paragraph.

7. **What are the potential benefits:** The anticipated benefits to children in your classroom include improving their early numeracy skills, which in turn increases their school readiness. The benefits for Head Start include the development of a practical and effective classroom activities targeted at improving children's early math skills.

8. **Points of contact:** Dr. James (Jim) DiPerna and Erin Reid are the lead researchers at Penn State University, and they can be reached via email (Erin – eer121@psu.edu; Jim - jdiperna@psu.edu) or phone (Erin – 814-863-9386; Jim - 814-863-2405) should you have any questions or comments about the project.

Appendix E

Teacher Consent Form

Informed Consent Form for Social Science Research
Teacher Consent Form
The Pennsylvania State University

Title of Project: **Number*Fun*:** Promoting Early Numeracy Skill Growth in Head Start Children

Principal Investigator: James C. DiPerna
105 CEDAR Building
University Park, PA 16802
(814) 863-2405; jdiperna@psu.edu

Other Investigator(s): Erin E. Reid

9. **Purpose of the Study:** The purpose of this research study is to develop and evaluate classroom activities intended to increase early numeracy skills in children attending Head Start.
10. **Procedures to be followed:** You will be asked to provide information about yourself such as your educational background and years of experience. You will be asked to keep a daily log of the early mathematics activities you lead in the classroom. You also will be asked to complete surveys on the learning and behavior of individual children in your classroom. If you are selected to implement the intervention in your classroom, you also will be asked to participate in the development of early numeracy instructional activities by reviewing materials and sharing your feedback with the investigators. Once the intervention has been developed, you will be trained to implement the program in your classroom. Trained research staff will periodically observe how you implement the program throughout the course of the study.
11. **Duration/Time:** This project will begin in January of 2008 and last until May 2009. You will spend about 20 hours per year in activities related to this study outside of the classroom (e.g., completing intervention logs and questionnaires). Teachers who are selected to implement the intervention in their classrooms will spend an additional 10 hours during Year 1 and 20 hours during Year 2 in project related activities (e.g., participating in training sessions and meeting with the project investigators).
12. **Statement of Confidentiality:** Your participation in this research is confidential. The data will be stored and secured at the Pennsylvania State University in a password-protected file. Only the persons in charge of the study will see the information collected through the study. No personally identifying information will be included in any published report. To guarantee confidentiality, you and participating children from your classroom will be assigned unique code numbers that will be placed on all forms and questionnaires instead of names. In the event of a publication or presentation resulting from the research, no personally identifiable information will be shared.
13. **Right to Ask Questions:** Please contact Erin Reid at (814) 863-9386 or James DiPerna at (814) 863-2405 with questions or concerns about this study.
14. **Payment for participation:** You will be compensated $25 per hour for time spent outside of the classroom on activities related to this project. Total payments within one calendar year that exceed $600 will require the University to report these payments to the IRS annually.

Appendix E (continued)

This may require you to claim the compensation that you receive for participation in this study as taxable income.

15. **Voluntary Participation:** Your decision to be in this research is voluntary. You can stop at any time. You do not have to answer any questions you do not want to answer. Refusal to

    take part in or withdrawing from this study will involve no penalty or loss of benefits you would receive otherwise.

You must be 18 years of age or older to consent to take part in this research study. If you agree to take part in this research study and the information outlined above, please sign your name and indicate the date below.
You will be given a copy of this form for your records.

______________________________________ ____________________
Participant Name (Please Print) Date

______________________________________ ____________________
Participant Signature Date

______________________________________ ____________________
Person Obtaining Consent Date

Appendix F

Parent Recruitment Letter: English

September 29, 2008

Dear Head Start & Pre K Counts Family,

Lebanon County Head Start and Pre K Counts are excited to tell you that we have recently entered into a 2-year partnership with Penn State University. The purpose of this partnership is to develop a classroom program that will improve the early number skills of the children. We look forward to this partnership because it will help our teachers to better serve children enrolled in Head Start and Pre K Counts. Attached to this page is a consent letter describing the purpose of the project. The letter explains how your child would be involved. It includes contact information for key project personnel, if you have any questions.

If you **want** your child to participate in the research partnership, **you do not need to do anything**.

If you **do not want** your child to participate in the project, **please complete the second page** of the letter and return it within 10 days of receiving this letter. (A return envelope is included if you decide that you do not want your child to participate in this project.)

We believe this partnership will benefit all children enrolled in Lebanon County Head Start and Pre K Counts. Please contact Erin Reid, the principal investigator, at 814-863-9386, if you have any questions or concerns about the project.

Thank you for your consideration of this opportunity and continued support of Head Start and Pre K Counts!

Sincerely,

Lebanon County Head Start and Pre K Counts Leadership Team

Appendix G

Parent Recruitment Letter: Spanish

Septembre 29, 2008

Estimada Familia de Head Start & Pre K Counts,

Lebanon County Head Start and Pre K Counts son emocionados a decirle que hemos entrado recientemente en una asociación de 2 años con Universidad de Penn State. El propósito de esta asociación es de desarrollar un programa de clase que mejorará las habilidades básico de aritméticas de los niños. Esperamos esta asociación ayudará a nuestros maestros a servir mejor a niños matriculados en Head Start y Pre K Counts. Ajunto a esta página es una carta del consentimiento que describe el propósito del proyecto. La carta explica cómo su niño sería implicado. Incluido es información de contacto para el personal clave del proyecto, si usted tiene cualquiera pregunta.

Si usted **quiere** que su niño tome parte en la sociedad de la investigación, **usted no necesita hacer nada.**

Si usted **no quiere** que su niño tome parte en el proyecto, completa **por favor la segunda página de la carta** y lo envías dentro de 10 días de recibir esta carta. (Un sobre de regreso es incluido si decide que usted no quiere que su niño tome parte en este proyecto).

Creemos que esta asociación beneficiará a todos niños matriculados en Head Start del Condado de Lebanon y Pre K Counts. Por favor haga contacto con Erin Reid, la investigadora principal, en 814-863-1939, si usted tiene cualquiera pregunta o preocupación sobre el proyecto.

¡Gracias para su consideración de esta oportunidad y apoyo continuado de Head Start y Pre K Counts!

Sinceramente,

El Equipo de Liderazgo de Head Start del Condado de Lebanon y Pre K Counts

Appendix H

Parent Consent Form: English

Opt-out Consent Form for Social Science Research
The Pennsylvania State University

Title of Project: Number*Fun*: Promoting Early Numeracy Skill Growth in Head Start Children

Principal Investigator:
James C. DiPerna
105 CEDAR Building
University Park, PA 16802
(814) 863-2405; jdiperna@psu.edu

Other Investigator(s): Erin E. Reid

1. **Purpose of the Study:** The purpose of this research study is to develop and evaluate classroom activities intended to increase Head Start children's understanding of number.
2. **Procedures to be followed:** In order to determine the effectiveness of the activities, your child will be asked to complete several measures of early mathematics skills throughout the school year. Trained research staff will visit your child's Head Start classroom and administer these tasks to your child in a quiet location in or near the classroom. At the end of each session, your child will receive stickers for participating. In addition, your child's classroom teacher will be asked to complete a form about your child's learning in the classroom.
3. **Duration/Time:** Research staff will work with your child five times throughout the school year. Each session will last for no more than 20 minutes. The total time your child will be asked to complete these measures is about 2 hours for the whole year.
4. **Statement of Confidentiality:** Your child's participation is confidential. Only the persons in charge of the study will see the information collected through the study. No personally identifying information will be included in any published report. To guarantee confidentiality, your child will be given a unique code number that will be placed on all forms and questionnaires instead of your child's name.
5. **Right to Ask Questions:** Please contact Erin Reid at (814) 863-9386 or James DiPerna at (814) 863-2405 with questions or concerns about this study.
6. **Voluntary Participation:** Your decision to allow your child to be tested by research staff and for your child's teacher to provide information about your child's learning is voluntary. You or your child can stop at any time and do not have to answer any questions you or your child do not want.

You must be 18 years of age or older to have your child take part in this research study.

You do not have to do anything for your child to participate in this study. Your child will automatically be enrolled in the study, unless you request otherwise.

Please keep this form for your records for future reference.

Appendix H (continued)

If you **DO NOT** want your child to participate in this study, please sign the form below and return it in the enclosed envelope to: Erin Reid, 226 CEDAR Building, University Park, PA 16802. A copy of this form will be returned to you.

______________________________ ______________________

Parent's Signature Date

____

Parent's name (please print)

____

Child's name (please print)

Appendix I

Parent Consent Form: Spanish

La Forma de Opción de no Participar en la Investigación de Ciencias Sociales Pennsylvania State University

Titulo del Proyectó: La Diversión de Numero: Promover el Crecimiento Temprano de la Habilidad de la conocimiento básico de aritmética en Niños de Head Start.

Investigador Principal: James C. DiPerna
105 CEDAR Building
University Park, PA 16802
(814) 863-2405; jdiperna@psu.edu

Otro Investigador(es): Erin E. Reid

1. **El Propósito del Estudio:** El propósito de este estudio de investigación es de desarrollar y evaluar las actividades con intención a aumentar la conocimiento básico de aritmética de niños de Head Start.
2. **Los Procedimientos Siguiente:** Para determinar la eficacia de las actividades, su niño será pedido completar varias medidas de habilidades tempranas de matemáticas durante todo el año escolar. Los personal entrenado de las investigación visitará su clase de niño y administrará estas tareas a su niño en una ubicación tranquila cerca del la clase. A finales de cada sesión, su niño recibirá pegatinas para participar. Además, la maestra será pedida completar una forma del aprendizaje de su niño en la clase.
3. **El Duración/Tiempo:** El Personal de Investigación trabajará con su niño cinco veces a través del año escolar. Cada sesión durará para no más de 20 minutos. El tiempo total que su niño será pedido completar estas medidas están acerca de 2 horas para el año entero.
4. **La Declaración de Confidencialidad:** Su participación del niño es confidencial. Sólo los dirigentes del estudio verán la información completa por el estudio. Ninguna información personalmente de identificación será incluida en cualquiera informe publicado. Para garantizar confidencialidad, su niño será dado un número asignando que será colocado un lugar de su nombre en todas formas.
5. **El Derecho de Hacer Preguntas:** Contacte por favor Erin Reid en (814) 863-9386 o James DiPerna en (814) 863-2405 con preguntas o preocupas de este estudio.
6. **La Participación Voluntaria:** Su decisión de permitir a su niño para ser probado por el personal investigador y para su maestro de niño para proporcionar información del aprendizaje de su niño es voluntario. Usted o su niño puede parar en cualquier tiempo y no tener que contestar ninguna pregunta usted o su niño no quiere.

Usted debe ser 18 años de la edad o más para que su niño participe en este estudio de investigación.

Usted no tiene que hacer nada para su niño participe este estudio. Su niño será matriculado automáticamente en el estudio, a menos que usted solicite de lo contrario.

Appendix I (continued)

Mantenga por favor esta forma para sus documentos para la referencia en el futuro.
Si usted no quiere que su niño tome parte en este estudio, firma por favor la forma debajo y lo vuelve en el sobre encerrado a Erin Reid, 226 CEDAR Building, University Park, PA 16802. Una copia de esta forma será regresado a usted.

____ ______________________

Firma de Padre Fecha

____

El nombre del padre (por favor imprima)

____

El Nombre de Niño (por favor imprima)

Appendix J

Skills Targeted by NumberFun Intervention

Domain	Skill
Number and Operations	Counting
	Recite number words in sequence (1-30)
	Subitize sets of up to 6
	Count with one-to-one correspondence (1-20)
	Count on from 2-19
	Name number that occurs before/after (1-20)
	Number knowledge (1-20)
	Cardinality
	Number recognition (0-20)
	Quantity discrimination (1-10)
	Estimation (1-20)
	Comparing and Ordering
	Quantity Comparison (more, less, equal)
	Ordinality (1-5)
	Order sets from least to most (1-10)
	Order written numbers (1-10)
	Adding To/Taking Away
	Adding/Subtracting by 1-4 with manipulatives
	Adding/Subtracting by 1-4 with number line
	Composing/Decomposing/Equal Partitioning
	Writing Numerals (0-10)
Patterns and Measurement	Patterns
	Copy/Extend/Create AB-AB pattern
	Copy/Extend/Create ABC-ABC pattern
	Copy/Extend/Create AA-BB pattern
	Copy/Extend/Create ABCD-ABCD pattern

(table continues)

(table continued)

Domain	Skill
Patterns and Measurement	Measurement
	Compare objects based on dimensions
	Sort/order/match based on dimensions
	Use non-standard measurement
	Data Analysis
	Gather information about self and others
	Create graph
	Interpret graphs
Geometry and Spatial Sense	Shape
	Identify basic shapes by name
	Identify basic shapes by physical properties
	Shape matching
	Compose/decompose shapes
	Location and Directions

Appendix K

Numbers & Operations

Number Wheel (1-10)

Children engage in gross motor activities as they count in sequence.

Setting:
Large group

Materials:
- Number wheel (1-10)

Advanced Preparation:
None

Prerequisite Skills:
Children must have mobility. For children who have physical disabilities, modify the activities to include motions they can perform (e.g., eye blinks, tongue thrusts, etc.).

Related Standards:
HS 3.1.3; 3.1.4
NCTM 1.1

Objectives:
- Children will count to 10 with adult assistance.
- Children will demonstrate one-to-one correspondence by counting and performing a physical movement at the same time.

Steps:
1. Introduce the activity. Say something like, **"We are going to practice counting. Each one of you is going to take a turn spinning this wheel. Whichever number the arrow lands on is the number of times we will move our bodies."**
2. Point to each number on the wheel and have children state its name.
3. Demonstrate the activity. Spin the wheel, call out the number, and instruct the children to clap that number of times. Begin clapping and start counting at a slow pace (approximately 1 second between claps). The rate of clapping can be increased as the children become more fluent counters. Pause and wait for the children to count with you if they do not automatically do so. Count each clap as it is being completed.
4. Give each child a turn spinning the wheel and calling out a movement. Repeat the procedures in Step 3 as needed.
5. Vary the loudness (and speed for more fluent counters) of the counting. That is, have the children whisper for some activities and speak loudly for others.

Downward Extension:
- Instead of the number wheel, use a die to help children become more fluent with counting 1-6.

Upward Extension:
- Have each child take turns spinning the wheel, drawing a card, and performing the activity by him/herself.
- Use the 10-20 number wheel
- Have the children identify the number to which the arrow is pointing.
- During one round near the end of the activity, do not join the children in counting. Listen to see if they are counting in sequence. If the group gets out of sequence, begin to count with them.

Appendix L

Geometry & Spatial Sense

Shape Sorting

Children will identify the basic shapes (circle, triangle, and square) and use terms such as "sides" to discuss the shapes.

Setting:
Large or small group

Materials:
- large shape cut-outs (circle, triangle, and square)
- variety of paper, felt, or foam shape cut-outs (different sizes and colors of circles, triangles, and squares)
- bowl or paper bag to hold the cut-outs
- poster board, easel with paper, or felt board
- tape

Advanced Preparation:
None

Prerequisite Skills:
Children should have experience counting to 4.

Related Standards:
HS 3.2.1, 3.2.4, 3.3.2, 3.3.3
NCTM 2.1, 2.2

Objectives:
- Children will identify basic geometric shapes (circle, triangle, and square).
- Children will identify the number of sides of the basic shapes.
- Children will sort shapes of different sizes into the appropriate category.

Steps:
1. Introduce the square. Say something like, **"Who can tell what the name of this shape is?"** pause for a response. **"This shape is a square. We know this is a square because it has 4 sides."** Count the number of sides with the children. **"How many sides does the square have?"** pause for a response**.**
2. Introduce the triangle. Say something like, **"Who can tell me what the name of this shape is?"** pause for a response. **"This shape is a triangle. We know this is a triangle because it has 3 sides."** Count the number of sides with the children. **"How many sides does the triangle have?"** pause for a response**.**
3. Introduce the circle. Say something like, **"Who can tell me the name of this shape?"** pause for a response. **"This shape is a circle. It is round. How many sides does a circle have?"** pause for a response. **"A circle doesn't have any sides.**
4. Attach the rectangle, square, triangle, and circle across the top of the poster board.
5. Introduce the activity. Hold up the bag containing the shape cutouts. **"Each of you is going to take a turn reaching into this bag and pulling out a shape. Then you will put your shape under its matching shape."**
6. Demonstrate the activity.
7. Allow one child at a time to choose a shape, name it, and place it on the poster board. Continue the activity until all of the shapes are gone.

Upward Extension:
- Introduce the concept of angles and have the children count the number angles for each shape.

Appendix M

Patterns & Measurement

Moving Patterns

Children create patterns using body movements.

Setting:
Large or small group

Materials:
None

Advanced Preparation:
None

Prerequisite Skills:
If some children in the classroom have limited mobility, include movements that they can perform (e.g., eye blinking, tongue thrusts, etc.).

Related Standards:
HS 3.3.1
NCTM 5.2

Objectives:
- Children will create and extend simple patterns (ABAB; ABCABC; AABB)

Steps:
1. Introduce the activity. **"Watch me."** Clap your hands twice and stomp your right foot twice. Repeat. **"Who can tell me what pattern I made?"** pause for a response. **"I clapped two times,"** clap twice, **"and I stomped my foot two times,"** stomp your right foot twice**. "Let's do it together."** Repeat the pattern. After stomping your foot twice say, **"What comes next?"** Repeat pattern a few times then add two hops to the end of the pattern. Clap hands, stomp feet, and ask **"What comes next?"** pause for a response. **"That's right, two hops. Let's do it together."** Repeat the pattern a few times.
2. Start a new pattern. Let the children decide the pattern using movements of their choice. If children tend to repeat the same type of patterns (e.g., AABB) encourage the children to try new things.
3. Continue the activity until several different pattern types have been created and mastered by most of the children.

Downward Extension:
- If the children are having difficulty creating patterns, lead the patterns for them.
- Name the movement as you act out the pattern.

Upward Extension:
- Encourage the children to create more complex patterns (e.g., ABCCABCC).

Appendix N

Internal Consistency and Item Analysis for EARLI Numeracy Probes.

	Counting Aloud	Counting Objects	Grouping	Measurement	Number Naming	Pattern Recognition
			Pre-test			
Age 3						
Alpha	.99	.83	.90	.83	.83	.87
p-value	.00-1.00	.28-.93	---	.36-71	.07-.76	.43-.64
Item-total *r*	.00-.77	.30-74	.51-.81	.07-.73	.30-.64	.02-.57
Age4						
Alpha	.99	.84	.94	.77	.87	.79
p-value	.00-1.00	.47-.69	---	.28-.97	.42-.94	.44-.81
Item-total *r*	.00-.78	.60-.85	.45-.89	.22-.76	.28-.79	.27-.77
			Post-test			
Age 3						
Alpha	.99	.88	.91	.79	.82	.89
p-value	.00-.97	.43-.95	---	.22-.86	.30-.84	.32-.65
Item-total *r*	.00-.86	.49-.73	.53-.83	-.14-.63	.23-.80	.25-.68
Age 4						
Alpha	.97	.88	.96	.82	.83	.72
p-value	.07-1.00	.57-.83	---	.47-1.00	.27-.77	.62-.83
Item-total *r*	.00-.85	.44-.83	.71-.93	.24-.75	.42-.79	.40-.75

Appendix O

Classroom Observation of Early Mathematics Instruction

Observer ID:	_______	Classroom ID:	_______	Date:	_______
Number of Children:	_______	Number of Adults:	_______	Setting:	_______________________
Number*Fun* Teachers Only		Activity Number:		Start Time _______	End Time _______

	Strongly Disagree	*Somewhat Disagree*	*Somewhat Agree*	*Strongly Agree*	*Not Observed*
Organization					
Appeared prepared for the activity	1	2	3	4	0
Had materials readily available	1	2	3	4	0
Pace of activity was appropriate	1	2	3	4	0
Managed the activity appropriately	1	2	3	4	0
Activity objectives were apparent	1	2	3	4	0
Engagement					
Conducted the activity with enthusiasm	1	2	3	4	0
Attempted to engage children in activity	1	2	3	4	0
Actively involved the children	1	2	3	4	0
Promoted idea that math is fun	1	2	3	4	0
Quality Instruction					0
Had high, realistic expectations	1	2	3	4	0
Adapted activities/tasks to child's ability	1	2	3	4	0
Acknowledged children's effort	1	2	3	4	0
Provided several opportunities for practice	1	2	3	4	0
Provided immediate feedback	1	2	3	4	0
Checked for understanding	1	2	3	4	0
Corrected misunderstanding	1	2	3	4	0

Appendix O (continued)

Promotion of Mathematical Thinking					
Engaged children's math thinking	1	2	3	4	0
Encouraged children to share ideas	1	2	3	4	0
Elaborated children's ideas	1	2	3	4	0
Encouraged reflection	1	2	3	4	0
Encouraged evaluation other's answers	1	2	3	4	0
Uses examples and non-examples	1	2	3	4	0
For Number*Fun* Teachers Only					
On schedule	1	2	3	4	0
Conducted activity as presented in manual	1	2	3	4	0
Significantly modified activity	1	2	3	4	0
Met activity objectives	1	2	3	4	0

Appendix P

Early Mathematics Daily Activity Log

Week of:

	Monday	Tuesday	Wednesday	Thursday	Friday
Activity #1					
Duration in minutes					
Brief Description					
Learning Objectives					
Activity #2					
Duration in minutes					
Brief Description					
Learning Objectives					
Activity #3					
Duration in minutes					
Brief Description					
Learning Objectives					

Appendix Q

Number*Fun* Daily Activity Log

Week of:

	Monday	Tuesday	Wednesday	Thursday	Friday
Number*Fun* Activity #1					
Activity #					
Duration in minutes					
Ease of implementation	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult
Difficulty level for majority of children	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult
Child engagement level	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much
Activity extensions used	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward
Was the activity modified?	Yes No	Yes No	Yes No	Yes No	Yes No
Describe Modifications					
Number*Fun* Activity #2					
Activity #					
Duration in minutes					
Ease of implementation	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult
Difficulty level for majority of children	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult
Child engagement level	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much
Activity extensions used	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward
Was the activity modified?	Yes No	Yes No	Yes No	Yes No	Yes No
Describe Modifications					

Appendix Q (continued)

Number*Fun* Daily Activity Log (continued)

	Monday	Tuesday	Wednesday	Thursday	Friday
Number*Fun* Activity #3					
Activity #					
Duration in minutes					
Ease of implementation	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult	Easy Moderate Difficult
Difficulty level for majority of children	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult	Too Easy Age-appropriate Difficult
Child engagement level	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much	Not at all A little Very Much
Activity extensions used	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward	None Downward Upward
Was the activity modified?	Yes No	Yes No	Yes No	Yes No	Yes No
Describe Modifications					
Non-Number*Fun* Activity #1					
Duration in minutes					
Brief Description					
Learning Objectives					
Non-Number*Fun* Activity #2					
Duration in minutes					
Brief Description					
Learning Objectives					

Appendix R

Number*Fun* Intervention Program
Teacher Feedback

Thank you for the time and effort you devoted to this project. In order to further improve this program, it would be appreciated if you would complete the following survey that asks questions related to your experiences while using the NumberFun Intervention Program. Please return the completed survey to Erin Reid by ***MM/DD/YY*** *in the envelope provided.*

Teacher ID: ________________ Date: ________________

Please circle the number that corresponds with how much you agree with the statement.

	Strongly Disagree	*Somewhat Disagree*	*Somewhat Agree*	*Strongly Agree*
Number*Fun* Manual				
The activities were presented in a clear and concise manner.	1	2	3	4
Providing more information about the theory and research behind the activities used in Number*Fun* would be helpful.	1	2	3	4
The activities were presented in a logical sequence.	1	2	3	4
If a substitute had to conduct an activity, she or he would have difficult following the manual.	1	2	3	4
The manual did not provide enough information to conduct the activities.	1	2	3	4
Number*Fun* Implementation				
Overall, the intervention activities were easy to implement.	1	2	3	4
A reasonable amount of time was required to prepare for the activities.	1	2	3	4
More professional development is needed to understand and implement the intervention.	1	2	3	4
A substitute teacher would have difficulty implementing the intervention as designed.	1	2	3	4
Number*Fun* Activities				
The children in my classroom enjoyed the activities.	1	2	3	4
I understood the concepts being taught.	1	2	3	4

Appendix R (continued)

Overall, the activities were developmentally appropriate.	1	2	3	4
The intervention activities encouraged children to use higher-order thinking skills.	1	2	3	4
The intervention materials were engaging.	1	2	3	4
The upward and downward extensions were useful in adapting the activities to meet the needs of individual children.	1	2	3	4
The concepts presented in the activities were too sophisticated for the children in my classroom.	1	2	3	4
The activities covered all of the important mathematical skill areas.	1	2	3	4
Overall, the activities were unique.	1	2	3	4
I modified many intervention activities to make them more appropriate for my children.	1	2	3	4
If the materials for each activity were not included, I would not use the intervention in the future.	1	2	3	4
The intervention materials were age appropriate.	1	2	3	4
The activities were too long in duration for the children in my classroom.	1	2	3	4
Satisfaction with Number*Fun*				
I would consider using this intervention in the future.	1	2	3	4
I would recommend using this intervention to other Head Start teachers.	1	2	3	4
I am pleased with the Number*Fun* intervention.	1	2	3	4

Appendix S

Number*Fun* Assessment Checklist

Date: ____________ Time: ________

Data Collector ID: ____________

Observer ID: ____________ Measure Observed: _________

	Observed	Not Observed
Accurately completed demographic information on protocol prior to seeing the child.		
Asked child for assent prior to testing.		
Examiner made efforts to establish rapport prior to testing.		
Testing materials were readily accessible but materials not in use are out of child's sight.		
Correct protocol and materials were used based on time point (fall, winter, or spring) and child's age.		
Began at correct starting point (EARLI probes – Item 1; TEMA-3 – Item 1 [Age 3], Item 7 [Age 4], or Item 15 [age 5]).		
All directions and test items were read verbatim.		
Followed basal rule for TEMA-3 (if 5 items in a row are not passed from the beginning point, examiner tested backward until 5 scores of 1 were obtained or until Item 1 was administered. Then continued testing).		
Repeated directions and/or test items upon request or if it was apparent child did not hear or understand the first time.		
Encouraged response when child did not respond.		
Unless stated otherwise in the manual, did not provide corrective feedback.		
Praised effort, good behavior, etc. but not performance.		
Administered all of the items for the EARLI probes.		
Followed ceiling rule for TEMA-3 only (discontinued testing after 5 consecutive scores of 0).		
Provided the child with reinforcement at the completion of every subtest (EARLI) or at the completion of the measure (TEMA-3)		

Appendix T

Descriptive Statistics for COEMI Items Averaged Across Activities

	Intervention						Comparison			
	1		2		3		4		5	
Subscales/Items	M	SD	M	SD	M	SD	M	SD	M	SD
Number of Activities	7		5		6		7		6	
Organization										
Appeared prepared for the activity	3.86	.38	4.00	0	4.00	0	4.00	0	4.00	0
Had materials readily available	4.00	0	4.00	0	3.83	.41	4.00	0	4.00	0
Pace of activity was appropriate	4.00	0	4.00	0	4.00	0	3.71	.49	3.67	.52
Managed the activity appropriately	3.86	.38	4.00	0	4.00	0	3.86	.38	3.67	.52
Activity objectives were apparent	4.00	0	4.00	0	4.00	0	3.86	.38	3.50	.55
Engagement										
Conducted the activity with enthusiasm	2.86	.38	4.00	0	3.50	.55	2.71	.76	2.50	.84
Attempted to engage children in activity	3.29	.49	4.00	0	4.00	0	3.00	.82	3.00	0
Actively involved the children	3.86	.38	4.00	0	4.00	0	3.14	.38	3.33	.52
Promoted idea that math is fun	3.00	.58	3.80	.45	3.50	.55	2.43	.79	1.83	.75
Quality										
Had high, realistic expectations	3.71	.49	4.00	0	3.67	.52	3.71	.49	3.17	.75
Adapted activities/tasks to child’s ability	2.86	1.35	4.00	0	3.50	.55	3.29	.95	3.00	0
Acknowledged children’s effort	3.00	1.16	3.80	.45	4.00	0	3.29	.95	2.83	.75
Provided several opportunities for practice	1.57	1.72	2.20	2.05	4.00	0	1.43	1.40	1.00	.89
Provided immediate feedback	3.43	.78	4.00	0	3.67	.52	3.86	.38	3.50	.84

(table continues)

(table continued)

	Intervention						Comparison			
	1		2		3		4		5	
Subscales/Items	M	SD	M	SD	M	SD	M	SD	M	SD
Checked for understanding	2.00	1.63	2.20	1.30	3.00	.63	2.14	1.35	1.00	.63
Corrected misunderstanding	1.71	2.14	3.80	.45	3.67	.52	2.57	1.40	2.33	1.86
Promotion of Mathematical Thinking										
Engaged children's math thinking	3.29	.76	3.80	.45	3.67	.52	3.14	.69	2.17	.75
Encouraged children to share ideas	3.00	1.00	2.60	1.52	2.33	.82	2.71	.95	1.67	1.03
Elaborated children's ideas	2.14	1.22	2.40	1.34	1.83	1.17	2.43	1.27	1.17	.98
Encouraged reflection	1.71	1.25	2.40	1.52	2.00	1.10	1.71	.76	.83	.41
Encouraged evaluation other's answers	2.86	1.68	3.80	.45	3.33	1.21	1.71	1.25	1.17	.98
Uses examples and non-examples	1.29	1.38	2.00	1.41	2.67	1.75	1.00	.58	.83	.41
Number*Fun* Fidelity										
On schedule	4.00[a]	0	3.40	.55	1.67	1.03	---	---	---	---
Conducted activity as presented in manual	4.00[a]	0	4.00	0	3.83	.41	---	---	---	---
Significantly modified activity[b]	4.00[a]	0	4.00	0	3.83	.41	---	---	---	---
Met activity objectives	4.00[a]	0	4.00	0	3.83	.41	---	---	---	---

Note. [a]Based on three activities. [b]Reverse coded.

Appendix U

Teacher Ratings of Satisfaction with NumberFun Curriculum

	Teacher Ratings			Total	
Domain / Items	Teacher 1	Teacher 2	Teacher 3	*M*	*SD*
Number*Fun* Manual					
The activities were presented in a clear and concise manner.	3	4	4	3.67	.58
Providing more information about the theory and research behind the activities used in Number*Fun* would be helpful.[a]	3	3	4	3.33	.58
The activities were presented in a logical sequence.	4	4	3	3.67	.58
If a substitute had to conduct an activity, she or he would have difficult following the manual. [a]	3	4	2	3.00	1.00
The manual did not provide enough information to conduct the activities. [a]	4	4	4	4.00	.00
Number*Fun* Implementation					
Overall, the intervention activities were easy to implement.	4	4	4	4.00	.00
A reasonable amount of time was required to prepare for the activities.	4	4	4	4.00	.00
More professional development is needed to understand and implement the intervention. [a]	2	4	4	3.33	1.16
A substitute teacher would have difficulty implementing the intervention as designed. [a]	3	4	3	3.33	.58

(table continues)

(table continued)

Domain / Items	Teacher Ratings			Total	
	Teacher 1	Teacher 2	Teacher 3	*M*	*SD*
Number*Fun* Activities					
The children in my classroom enjoyed the activities.	4	4	4	4.00	.00
I understood the concepts being taught.	4	4	4	4.00	.00
Overall, the activities were developmentally appropriate.	4	4	4	4.00	.00
The intervention activities encouraged children to use higher-order thinking skills.	3	4	4	3.67	.58
The intervention materials were engaging.	4	4	4	4.00	.00
The upward and downward extensions were useful in adapting the activities to meet the needs of individual children.	3	3	4	3.33	.58
The concepts presented in the activities were too sophisticated for the children in my classroom. [a]	3	3	4	3.33	.58
The activities covered all of the important mathematical skill areas.	4	4	4	4.00	.00
Overall, the activities were unique.	3	4	3	3.33	.58
I modified many intervention activities to make them more appropriate for my children.[a]	2	3	4	3.00	1.00
If the materials for each activity were not included, I would not use the intervention in the future. [a]	3	4	3	3.33	.58
The intervention materials were age appropriate.	4	4	4	4.00	.00

(table continues)

(table continued)

Domain / Items	Teacher Ratings			Total	
	Teacher 1	Teacher 2	Teacher 3	*M*	*SD*
The activities were too long in duration for the children in my classroom.[a]	2	2	1	1.67	.58
Satisfaction with Number*Fun*					
I would consider using this intervention in the future.	4	4	4	4.00	.00
I would recommend using this intervention to other Head Start teachers.	4	4	4	4.00	.00
I am pleased with the Number*Fun* intervention.	4	4	4	4.00	.00

Note. [a]Reverse coded.

Appendix V

Descriptive Statistics for the Test of Early Mathematics Ability-3rd Edition Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	7.76	6.46	0.76	-0.51	11.89	8.92	0.96	0.81
1	7.69	6.04	0.73	-0.38	13.00	9.19	0.69	0.12
2	8.15	5.98	0.58	-0.49	11.61	8.18	0.97	1.22
3	7.82	5.97	0.66	-0.46	12.20	9.22	0.88	0.39
4	7.84	5.97	0.65	-0.46	12.24	9.06	0.76	0.18
5	8.39	6.41	0.69	-0.32	11.82	8.77	0.94	0.64
6	7.88	6.13	0.90	0.39	12.39	8.79	0.83	0.38
7	7.76	5.92	0.72	-0.29	11.64	8.72	0.90	0.59
8	7.89	6.22	0.69	-0.58	12.51	9.03	0.72	0.17
9	7.32	5.79	0.78	-0.06	11.67	8.56	0.87	0.62
10	8.44	6.46	0.77	-0.21	12.33	8.67	0.82	0.51
11	7.30	5.79	0.84	-0.04	13.09	9.18	0.74	-0.01
12	8.05	5.98	0.61	-0.45	11.76	8.40	0.97	0.85
13	8.24	6.30	0.82	-0.11	11.97	8.99	0.91	0.49
14	7.68	5.79	0.77	-0.11	12.98	8.86	0.62	0.02
15	7.64	5.63	0.76	0.03	11.71	8.98	0.95	0.69
16	7.92	5.91	0.85	0.10	12.38	8.90	0.66	0.11
17	7.67	6.27	0.85	-0.16	11.96	8.71	0.81	0.45
18	7.72	5.90	0.66	-0.40	11.72	8.45	0.85	0.70
19	7.56	6.05	0.73	-0.16	12.20	8.76	0.92	0.72
20	7.60	5.89	0.75	-0.26	11.69	8.68	0.80	0.47
21	8.32	6.39	0.81	-0.11	11.42	8.60	0.95	0.82
22	8.27	6.41	0.86	0.18	12.18	9.04	0.89	0.53
23	7.48	5.71	0.92	0.23	11.84	9.26	0.92	0.40
24	7.74	5.85	0.70	-0.33	11.78	8.88	0.94	0.73
25	7.72	5.89	0.76	-0.19	12.47	9.07	0.76	0.21
26	8.11	6.11	0.68	-0.44	12.19	8.51	0.80	0.53

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	7.70	5.95	0.76	-0.15	12.10	8.71	0.93	0.62
28	7.79	5.89	0.72	-0.32	12.36	8.81	0.75	0.28
29	7.35	6.11	0.89	-0.04	11.92	8.70	1.04	1.13
30	8.07	6.26	0.78	-0.05	12.26	8.67	0.87	0.54
31	7.52	5.74	0.70	-0.08	11.28	8.81	0.88	0.56
32	7.79	6.17	0.84	0.08	12.86	9.14	0.79	0.22
33	7.86	5.97	0.63	-0.38	13.05	8.49	0.66	0.28
34	7.72	6.38	0.96	0.31	11.58	8.41	1.00	1.04
35	7.95	6.30	0.82	-0.19	11.38	8.65	1.03	0.94
36	7.76	6.14	0.70	-0.43	12.10	8.36	0.91	0.90
37	7.82	5.81	0.75	-0.25	11.56	8.48	0.95	0.88
38	7.59	5.84	0.81	-0.24	12.36	9.38	0.77	0.03

Note. Raw scores are reported.
[a]For original data $N = 62$; for imputed datasets $N = 96$. [b]For original data $N = 71$; for imputed datasets $N = 96$.

Appendix W

Descriptive Statistics for the Counting Aloud EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	-0.56	0.64	1.14	2.09	-0.30	0.89	1.16	2.11
1	-0.48	0.76	0.60	-0.10	-0.18	1.08	0.61	0.14
2	-0.53	0.75	0.57	0.70	-0.26	0.89	0.82	0.85
3	-0.58	0.72	0.56	0.75	-0.21	1.02	0.61	0.72
4	-0.51	0.75	0.82	0.34	-0.19	0.99	0.69	0.55
5	-0.51	0.70	0.51	0.18	-0.30	0.87	0.90	1.58
6	-0.55	0.78	0.51	0.25	-0.14	0.98	0.60	0.22
7	-0.57	0.64	0.95	1.59	-0.40	1.05	0.54	0.86
8	-0.48	0.71	0.77	0.53	-0.12	1.07	0.41	0.35
9	-0.53	0.65	0.66	0.79	-0.26	1.03	0.64	0.85
10	-0.50	0.75	0.52	0.19	-0.45	0.89	0.81	1.66
11	-0.49	0.69	0.51	0.46	-0.22	0.97	0.93	0.91
12	-0.55	0.70	0.59	1.15	-0.27	0.98	0.48	0.86
13	-0.57	0.69	0.84	0.72	-0.33	0.83	1.10	2.00
14	-0.49	0.67	0.71	0.67	-0.38	0.99	0.55	0.79
15	-0.54	0.68	0.42	0.80	-0.30	0.97	0.69	1.14
16	-0.60	0.70	0.80	1.19	-0.19	0.96	0.57	0.57
17	-0.52	0.74	0.45	0.29	-0.31	0.97	0.79	0.91
18	-0.56	0.66	0.72	0.80	-0.26	0.88	0.70	1.11
19	-0.58	0.69	0.68	0.67	-0.33	1.01	0.41	1.22
20	-0.53	0.71	0.70	0.56	-0.33	1.03	0.59	1.05
21	-0.51	0.65	0.94	1.34	-0.29	0.94	0.82	1.14
22	-0.61	0.70	0.60	0.96	-0.28	0.87	0.70	1.29
23	-0.49	0.68	0.58	0.45	-0.33	0.97	0.44	0.81
24	-0.56	0.65	0.80	1.56	-0.29	0.94	0.75	0.78
25	-0.51	0.66	0.85	0.75	-0.32	0.97	0.61	0.62
26	-0.49	0.70	0.71	0.32	-0.20	0.91	0.60	1.11

(table continues)

(table continued)

Imputation	Pretest[a]				Post-test[b]			
	M	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	-0.53	0.70	0.70	0.58	-0.25	0.97	0.66	0.91
28	-0.58	0.67	0.77	0.98	-0.16	0.91	0.50	0.66
29	-0.55	0.70	0.73	0.92	-0.32	0.94	0.75	0.81
30	-0.55	0.68	0.70	0.54	-0.29	0.93	0.50	1.01
31	-0.54	0.71	0.59	0.31	-0.28	1.04	0.27	0.61
32	-0.50	0.69	0.74	0.68	-0.22	1.07	0.51	0.01
33	-0.56	0.69	0.97	1.43	-0.22	0.93	0.79	0.69
34	-0.56	0.67	0.76	1.10	-0.27	0.86	1.01	1.76
35	-0.53	0.72	0.77	0.54	-0.29	1.01	0.54	0.39
36	-0.49	0.69	0.52	0.24	-0.28	0.91	0.72	0.92
37	-0.52	0.71	0.84	0.75	-0.35	0.96	0.67	1.15
38	-0.57	0.67	0.95	1.01	-0.19	1.02	0.77	0.39

Note. IRT scaled scores are reported.
[a]For original data $N = 62$; for imputed datasets $N = 96$. [b]For original data $N = 65$; for imputed datasets $N = 96$.

Appendix X

Descriptive Statistics for the Counting Objects EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	-0.75	1.41	0.22	-0.68	-0.40	1.01	-0.71	0.37
1	-0.92	1.46	0.04	-0.30	-0.20	1.07	-0.40	0.40
2	-0.53	1.49	0.19	-0.40	-0.25	1.08	-0.42	0.40
3	-0.77	1.46	0.04	-0.29	-0.38	0.93	-0.66	0.59
4	-0.75	1.56	0.44	-0.20	-0.29	1.08	-0.36	0.12
5	-0.71	1.49	0.09	-0.60	-0.28	0.94	-0.73	0.86
6	-0.61	1.54	0.30	-0.28	-0.27	1.10	-0.33	0.33
7	-0.72	1.49	-0.01	-0.49	-0.31	1.05	-0.58	0.21
8	-0.86	1.47	0.11	-0.53	-0.25	1.08	-0.25	0.18
9	-0.80	1.46	0.07	-0.51	-0.33	1.01	-0.41	0.32
10	-0.66	1.49	0.26	-0.56	-0.43	1.16	-0.47	0.01
11	-0.82	1.44	0.05	-0.51	-0.32	1.02	-0.54	0.43
12	-0.73	1.63	0.19	-0.18	-0.27	1.07	-0.29	0.38
13	-0.78	1.38	0.04	-0.34	-0.25	1.08	-0.38	0.08
14	-0.68	1.35	0.19	-0.71	-0.45	1.02	-0.54	-0.11
15	-0.71	1.38	0.00	-0.51	-0.30	1.09	-0.23	0.29
16	-0.74	1.59	0.22	-0.30	-0.44	0.99	-0.55	0.02
17	-0.99	1.47	-0.08	-0.32	-0.33	1.04	-0.36	0.20
18	-0.88	1.43	0.11	-0.61	-0.38	1.05	-0.51	-0.23
19	-0.80	1.53	0.30	-0.01	-0.42	1.00	-0.34	0.35
20	-0.82	1.47	0.21	-0.48	-0.32	1.04	-0.55	0.34
21	-0.79	1.62	0.23	-0.53	-0.34	1.02	-0.43	0.24
22	-0.79	1.47	0.18	-0.38	-0.35	1.00	-0.40	0.33
23	-0.55	1.43	0.04	-0.63	-0.40	1.07	-0.43	-0.12
24	-0.67	1.54	0.02	-0.29	-0.27	0.98	-0.68	0.47
25	-0.97	1.44	-0.08	-0.20	-0.24	1.12	-0.55	0.31
26	-0.87	1.47	0.17	-0.62	-0.39	1.01	-0.56	0.11

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	-0.74	1.53	-0.01	0.22	-0.28	1.07	-0.44	0.23
28	-0.80	1.41	0.30	-0.04	-0.26	1.06	-0.56	-0.01
29	-0.78	1.46	0.08	-0.47	-0.27	0.98	-0.62	0.51
30	-0.65	1.42	0.29	-0.38	-0.32	1.03	-0.37	0.31
31	-0.83	1.65	-0.23	-0.25	-0.43	1.04	-0.18	0.29
32	-0.84	1.57	0.16	-0.62	-0.41	0.97	-0.63	0.35
33	-0.63	1.54	-0.04	-0.39	-0.29	1.11	-0.36	0.05
34	-0.70	1.54	0.15	-0.40	-0.29	1.04	-0.36	0.29
35	-0.70	1.48	0.24	-0.29	-0.32	1.09	-0.29	0.37
36	-0.88	1.68	0.18	-0.47	-0.30	0.98	-0.66	0.36
37	-0.73	1.46	0.31	-0.76	-0.42	0.99	-0.51	-0.07
38	-0.77	1.46	0.45	-0.17	-0.29	1.01	-0.55	0.29

Note. IRT scaled scores are reported.
[a]For original data N = 61; for imputed datasets N = 96. [b]For original data N = 67; for imputed datasets N = 96.

Appendix Y

Descriptive Statistics for the Grouping EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	0.73	0.92	0.18	-0.51	1.27	1.04	-0.13	-0.11
1	0.80	1.06	0.04	-0.38	1.47	1.11	-0.03	-0.08
2	0.68	0.93	0.03	-0.40	1.34	1.04	-0.12	-0.27
3	0.68	1.04	0.01	-0.33	1.34	1.13	0.00	-0.30
4	0.59	0.97	-0.20	-0.09	1.35	1.08	-0.03	-0.14
5	0.78	1.02	0.18	-0.07	1.34	1.07	0.05	-0.36
6	0.60	1.00	-0.15	-0.13	1.34	1.10	-0.02	-0.41
7	0.72	1.08	0.25	-0.49	1.22	1.24	0.12	-0.19
8	0.64	0.99	0.32	-0.25	1.26	1.17	-0.26	-0.22
9	0.70	0.95	0.17	-0.37	1.26	1.08	-0.17	-0.19
10	0.60	1.10	0.00	-0.38	1.19	1.08	-0.17	-0.46
11	0.60	1.02	-0.08	-0.64	1.39	1.12	0.01	0.09
12	0.79	0.95	0.09	-0.64	1.22	1.00	0.03	-0.25
13	0.64	0.96	0.17	-0.67	1.33	1.04	-0.03	-0.22
14	0.76	0.96	-0.12	-0.06	1.27	1.07	-0.22	-0.23
15	0.59	1.01	-0.02	-0.30	1.21	1.18	0.02	-0.11
16	0.72	0.94	0.01	-0.65	1.28	1.02	-0.05	-0.19
17	0.74	0.97	0.10	-0.59	1.21	1.07	-0.03	-0.17
18	0.70	0.93	0.36	0.05	1.22	1.11	-0.17	-0.21
19	0.61	1.06	-0.13	-0.36	1.25	1.15	0.09	0.04
20	0.72	0.98	0.01	-0.66	1.21	1.09	-0.11	-0.22
21	0.66	1.09	-0.25	-0.23	1.32	1.06	-0.06	-0.32
22	0.68	0.95	0.06	-0.70	1.30	1.05	-0.05	-0.02
23	0.57	0.97	0.01	-0.28	1.17	1.19	-0.16	-0.25
24	0.70	0.96	0.14	-0.21	1.35	0.98	-0.13	0.04
25	0.66	0.96	-0.09	-0.17	1.39	1.05	-0.04	-0.15
26	0.65	1.00	0.00	-0.10	1.37	1.06	-0.14	-0.40

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	0.62	1.08	0.20	-0.25	1.29	1.20	-0.20	-0.14
28	0.77	0.94	0.08	-0.83	1.40	1.10	-0.01	-0.19
29	0.63	1.04	-0.24	-0.29	1.34	0.99	-0.10	-0.06
30	0.65	1.07	0.07	-0.57	1.37	1.06	-0.05	-0.23
31	0.70	0.94	0.18	-0.64	1.20	1.05	0.15	-0.26
32	0.60	0.89	0.21	-0.37	1.23	1.07	0.05	-0.30
33	0.72	1.05	0.03	-0.28	1.31	1.12	0.08	-0.41
34	0.68	1.04	-0.01	-0.86	1.14	1.16	0.02	-0.31
35	0.74	1.05	-0.06	-0.51	1.24	1.10	0.06	-0.28
36	0.68	0.98	0.34	-0.07	1.34	1.15	-0.10	-0.37
37	0.60	1.07	-0.25	-0.25	1.11	1.17	-0.23	-0.06
38	0.68	1.00	-0.20	-0.32	1.35	1.10	-0.04	-0.32

Note. IRT scaled scores are reported.
[a]For original data N = 61; for imputed datasets N = 96. [b]For original data N = 65 ; for imputed datasets N = 96.

Appendix Z

Descriptive Statistics for the Measurement EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	-0.72	0.82	-0.23	-0.72	-0.45	1.03	-0.38	0.28
1	-0.81	0.96	-0.18	-0.37	-0.22	1.15	-0.07	0.21
2	-0.76	0.86	-0.04	-0.56	-0.50	1.10	-0.25	0.18
3	-0.65	0.89	-0.19	-0.55	-0.39	1.06	-0.23	0.07
4	-0.72	0.91	0.00	-0.48	-0.41	1.03	-0.05	0.20
5	-0.65	0.94	0.15	-0.45	-0.43	1.10	0.00	-0.06
6	-0.82	0.86	-0.11	-0.55	-0.38	1.18	-0.20	0.30
7	-0.75	0.87	0.02	-0.65	-0.32	1.07	-0.15	0.20
8	-0.78	0.91	0.00	-0.54	-0.33	1.07	-0.22	-0.04
9	-0.73	0.96	-0.27	-0.32	-0.42	1.01	-0.35	-0.04
10	-0.84	0.93	-0.43	-0.36	-0.49	1.02	-0.22	-0.20
11	-0.84	0.87	-0.27	-0.30	-0.36	1.03	-0.24	0.18
12	-0.76	0.88	-0.12	-0.67	-0.48	1.06	-0.04	-0.11
13	-0.71	0.87	0.13	-0.75	-0.37	1.00	-0.31	0.16
14	-0.76	0.86	-0.22	-0.51	-0.47	1.11	-0.33	-0.16
15	-0.74	0.90	-0.39	-0.30	-0.42	1.10	-0.12	-0.20
16	-0.74	0.93	-0.20	-0.45	-0.40	1.08	0.03	0.25
17	-0.70	0.91	-0.14	-0.21	-0.38	1.11	-0.19	0.33
18	-0.83	0.94	-0.11	-0.44	-0.40	1.10	-0.25	0.03
19	-0.72	0.92	-0.19	-0.05	-0.52	1.03	-0.17	0.15
20	-0.76	0.89	-0.04	-0.51	-0.40	1.16	-0.26	-0.01
21	-0.86	0.88	-0.35	-0.34	-0.41	1.03	-0.12	0.46
22	-0.84	0.95	-0.48	-0.28	-0.47	1.01	0.00	0.37
23	-0.84	0.82	-0.22	-0.48	-0.37	1.10	-0.03	0.28
24	-0.84	0.96	-0.15	-0.37	-0.34	1.08	-0.26	0.03
25	-0.71	0.94	0.17	-0.33	-0.29	1.18	-0.03	0.05
26	-0.78	1.02	-0.18	-0.49	-0.45	1.07	-0.15	0.27

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	-0.66	0.87	-0.14	-0.43	-0.40	1.05	-0.20	0.27
28	-0.75	0.91	-0.10	-0.42	-0.39	1.11	-0.20	-0.06
29	-0.65	0.95	0.06	-0.44	-0.35	1.07	-0.10	-0.01
30	-0.85	0.90	-0.10	-0.72	-0.43	1.04	-0.36	0.07
31	-0.76	0.90	-0.17	-0.71	-0.43	1.04	-0.08	0.53
32	-0.83	0.93	-0.39	-0.36	-0.43	1.06	-0.16	-0.24
33	-0.82	0.96	-0.28	-0.66	-0.28	1.13	-0.26	0.06
34	-0.62	0.91	-0.07	-0.45	-0.47	1.04	-0.34	0.02
35	-0.78	1.01	-0.08	-0.51	-0.49	1.04	-0.17	-0.06
36	-0.75	0.91	-0.38	-0.23	-0.45	1.08	-0.18	-0.18
37	-0.77	0.90	-0.14	-0.21	-0.59	1.00	-0.18	-0.07
38	-0.69	0.96	-0.13	-0.67	-0.31	1.15	-0.03	0.04

Note. IRT scaled scores are reported.
[a]For original data N = 60; for imputed datasets N = 96. [b]For original data N = 67; for imputed datasets N = 96.

Appendix AA

Descriptive Statistics for the Number Naming EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	-0.56	0.51	0.46	0.62	-0.32	0.86	1.40	2.16
1	-0.60	0.61	0.35	0.17	-0.32	1.00	0.53	0.73
2	-0.56	0.54	0.32	0.12	-0.31	0.87	0.89	1.00
3	-0.49	0.58	0.80	0.71	-0.38	0.88	1.19	1.84
4	-0.62	0.53	-0.10	0.75	-0.30	0.89	1.05	1.11
5	-0.56	0.56	0.11	0.44	-0.36	0.95	0.92	1.01
6	-0.59	0.54	0.12	0.17	-0.39	0.88	0.75	1.53
7	-0.52	0.61	0.53	0.03	-0.47	0.96	0.51	1.11
8	-0.54	0.56	0.14	0.20	-0.34	0.86	0.82	1.49
9	-0.62	0.60	0.24	0.26	-0.43	0.96	0.54	0.93
10	-0.55	0.49	0.36	0.34	-0.41	0.87	0.83	1.67
11	-0.61	0.54	0.11	0.41	-0.29	0.88	0.96	0.76
12	-0.65	0.54	0.42	0.40	-0.48	0.98	0.48	0.91
13	-0.57	0.56	0.02	0.21	-0.45	1.03	0.69	0.84
14	-0.59	0.59	0.07	0.24	-0.39	0.91	0.63	1.33
15	-0.63	0.56	0.22	0.21	-0.39	0.87	1.00	1.39
16	-0.60	0.54	0.29	0.05	-0.33	0.89	0.70	0.97
17	-0.63	0.55	0.35	0.38	-0.44	0.91	0.65	1.79
18	-0.57	0.57	0.57	0.55	-0.31	0.81	1.06	1.82
19	-0.55	0.53	0.28	0.45	-0.34	0.92	0.84	1.00
20	-0.60	0.56	0.22	0.19	-0.37	0.86	1.04	1.51
21	-0.56	0.54	0.68	0.68	-0.45	0.94	0.87	1.39
22	-0.59	0.55	0.02	0.31	-0.37	0.88	1.10	1.92
23	-0.63	0.56	0.06	0.00	-0.40	0.85	1.12	1.93
24	-0.61	0.56	0.45	0.70	-0.38	0.91	0.81	1.02
25	-0.57	0.53	0.47	0.61	-0.31	0.97	0.93	0.80
26	-0.59	0.50	0.27	0.41	-0.46	0.92	0.65	1.52

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	-0.57	0.53	0.26	0.39	-0.38	0.89	0.90	1.09
28	-0.53	0.51	0.34	0.01	-0.38	0.84	0.92	2.10
29	-0.64	0.56	0.28	0.32	-0.35	0.85	0.92	1.75
30	-0.63	0.59	0.15	-0.06	-0.31	0.89	0.68	1.01
31	-0.56	0.60	0.39	0.23	-0.41	0.91	0.55	1.39
32	-0.62	0.53	0.27	0.25	-0.31	0.86	0.89	1.36
33	-0.58	0.56	0.35	0.42	-0.35	0.85	1.08	1.45
34	-0.55	0.57	0.16	-0.10	-0.39	0.85	1.05	1.62
35	-0.61	0.49	0.46	0.93	-0.49	0.91	0.94	1.27
36	-0.58	0.56	0.31	0.31	-0.38	0.92	0.73	1.10
37	-0.56	0.54	0.25	0.04	-0.38	0.89	0.90	1.62
38	-0.63	0.59	0.14	0.19	-0.38	0.91	1.00	1.28

Note. IRT scaled scores are reported.
[a]For original data N = 62; for imputed datasets N = 96. [b]For original data N = 67; for imputed datasets N = 96.

Appendix AB

Descriptive Statistics for the Pattern Recognition EARLI Numeracy Probe Based on Original and Imputed Datasets

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
Original	-0.32	0.64	1.29	1.18	-0.34	0.97	-0.50	0.81
1	-0.35	0.71	0.82	0.45	-0.18	1.05	-0.27	0.55
2	-0.36	0.73	0.32	0.55	-0.30	0.99	-0.26	0.32
3	-0.36	0.73	0.62	0.19	-0.19	1.05	-0.23	0.36
4	-0.39	0.64	0.90	0.78	-0.24	1.00	-0.15	0.65
5	-0.31	0.69	0.58	0.12	-0.28	1.05	-0.30	-0.01
6	-0.37	0.72	0.53	0.18	-0.23	1.11	-0.19	0.21
7	-0.35	0.67	1.00	0.74	-0.35	1.02	-0.38	0.11
8	-0.34	0.77	0.38	0.20	-0.20	1.10	0.00	0.26
9	-0.35	0.70	0.41	0.33	-0.27	1.07	-0.58	0.57
10	-0.36	0.64	0.73	1.08	-0.31	1.12	-0.23	0.30
11	-0.33	0.69	0.80	0.07	-0.20	1.02	-0.23	0.71
12	-0.35	0.72	0.72	0.22	-0.30	0.96	-0.21	0.49
13	-0.41	0.67	0.83	0.80	-0.33	1.05	-0.23	0.04
14	-0.29	0.72	0.51	0.04	-0.22	1.19	-0.11	0.31
15	-0.37	0.68	0.68	0.52	-0.25	1.11	-0.19	0.10
16	-0.34	0.71	0.45	0.20	-0.29	0.98	-0.36	0.24
17	-0.40	0.72	0.31	0.65	-0.27	0.93	-0.30	0.88
18	-0.40	0.67	0.60	0.55	-0.30	1.01	-0.21	0.49
19	-0.35	0.73	0.62	0.24	-0.39	0.99	-0.31	0.43
20	-0.38	0.70	0.83	0.49	-0.29	1.15	-0.23	0.16
21	-0.31	0.72	0.48	-0.04	-0.34	1.02	-0.11	0.04
22	-0.41	0.68	0.88	0.78	-0.32	1.05	0.06	0.91
23	-0.38	0.79	0.36	0.43	-0.24	1.02	-0.28	0.15
24	-0.28	0.74	0.42	0.14	-0.22	1.00	-0.22	0.64
25	-0.41	0.69	0.71	0.44	-0.22	1.15	0.13	0.00
26	-0.28	0.71	0.82	-0.01	-0.21	1.10	-0.09	0.18

(table continues)

(table continued)

	Pretest[a]				Post-test[b]			
Imputation	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
27	-0.39	0.70	0.56	0.14	-0.25	1.05	-0.16	0.59
28	-0.37	0.68	0.74	0.80	-0.11	1.04	-0.22	0.71
29	-0.34	0.69	1.01	0.38	-0.25	1.05	-0.12	0.29
30	-0.31	0.71	0.74	0.45	-0.28	1.08	-0.25	0.50
31	-0.40	0.66	0.85	0.59	-0.32	1.13	-0.09	0.29
32	-0.31	0.66	0.53	0.46	-0.31	1.01	-0.37	0.24
33	-0.28	0.69	0.56	0.10	-0.26	1.11	-0.12	0.02
34	-0.33	0.71	0.47	0.40	-0.33	1.04	-0.45	0.15
35	-0.40	0.66	0.76	0.62	-0.24	1.10	-0.27	0.29
36	-0.40	0.68	0.77	0.43	-0.22	0.99	-0.26	0.55
37	-0.37	0.75	0.50	0.24	-0.28	1.06	-0.36	0.24
38	-0.34	0.70	0.88	0.26	-0.18	1.06	0.06	0.73

Note. IRT scaled scores are reported.
[a]For original data N = 60; for imputed datasets N = 96. [b]For original data N = 66; for imputed datasets N = 96.

Appendix AC

Pooled Results of Reversal of Multiple Regression Analyses for TEMA-3

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-2.88	1.99	-1.44	.15	.00 - .08	.00 - .08
Step 2						
Gender	2.76	1.76	1.57	.12		
Race	-1.24	2.56	-.48	.63		
Age	.68	.14	4.85	<.01	.21 - .37	.17 - .34
Step 3						
Pretest	.78	.17	4.63	< .01	.38 - .70	.11 - .40

Appendix AD

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Counting Aloud

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	.05	.25	.21	.84	.00 - .04	.00 - .04
Step 2						
Gender	.25	.22	1.15	.25		
Race	-.27	.35	-.76	.45		
Age	.05	.02	2.80	<.01	.10 - .23	.10 - .23
Step 3						
Pretest	.45	.20	2.31	.02	.12 - .34	.01 - .21

Appendix AE

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Counting Objects

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-.30	.25	-1.22	.22	.00 - .07	.00 - .07
Step 2						
Gender	.38	.23	1.64	.10		
Race	-.24	.32	-.75	.46		
Age	.05	.02	2.92	<.01	.10 - .28	.08 - .27
Step 3						
Pretest	.17	.10	1.83	.07	.12 - .36	.00 - .14

Appendix AF

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Grouping

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-.00	.27	-.01	.99	.00 - .03	.00 - .03
Step 2						
Gender	.38	.23	1.66	.10		
Race	-.44	.33	-1.32	.19		
Age	.08	.02	4.61	< .01	.20 - .41	.19 - .41
Step 3						
Pretest	.31	.15	2.16	.03	.27 - .52	.01 - .17

Appendix AG

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Measurement

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-.15	.26	-.56	.56	.00 - .04	.00 - .04
Step 2						
Gender	.25	.25	.99	.32		
Race	-.63	.36	-1.74	.08		
Age	.05	.02	2.71	<.01	.06 - .30	.06 - .28
Step 3						
Pretest	.32	.15	2.12	.04	.08 - .42	.00 - .14

Appendix AH

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Number Naming

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-.06	.21	-.27	.79	.00 - .02	.00 - .02
Step 2						
Gender	.07	.19	.37	.71		
Race	-.05	.29	-.18	.86		
Age	.06	.02	3.37	<.01	.09 - .21	.08 - .21
Step 3						
Pretest	.89	.24	3.75	< .01	.24 - .56	.08 - .38

Appendix AI

Pooled Results of Reversal of Multiple Regression Analyses for EARLI Pattern Recognition

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Condition	-.05	.27	-.17	.87	.00 - .04	.00 - .04
Step 2						
Gender	.21	.24	.88	.38		
Race	-.22	.35	-.62	.53		
Age	.05	.02	2.39	.02	.04 - .17	.04 - .17
Step 3						
Pretest	.47	.25	1.93	.06	.06 - .45	.00 - .31

Appendix AJ

Pooled Results of Multiple Regression Analyses for EARLI Counting Aloud with Classroom Added as a Predictor

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.56	.19	2.96	<.01	.04 - .30	.04 - .30
Step 2						
Age	.04	.02	2.14	.03	.12 - .34	.01 - .16
Step 3						
Teacher 2	.07	.23	.31	.75		
Teacher 3	-.50	.25	-.20	.84		
Teacher 1	.00	.30	.01	.99	.12 - .37	.00 - .06

Appendix AK

Pooled Results of Multiple Regression Analyses for EARLI Counting Objects with Classroom Added as a Predictor

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.24	.09	2.68	<.01	.03 - .26	.03 - .26
Step 2						
Age	.04	.02	2.36	.02	.09 - .30	.02 - .16
Step 3						
Teacher 2	-.06	.25	-.25	.80		
Teacher 3	-.23	.25	-.92	.36		
Teacher 1	.22	.31	.70	.49	.13 - .31	.00 - .10

Appendix AL

Pooled Results of Multiple Regression Analyses for EARLI Grouping with Classroom Added as a Predictor

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.49	.14	3.43	<.01	.09 - .32	.09 - .32
Step 2						
Age	.06	.02	3.14	<.02	.22 - .44	.03 - .21
Step 3						
Teacher 2	.04	.25	.14	.89		
Teacher 3	.28	.23	1.22	.22		
Teacher 1	-.29	.29	-1.00	.32	.24 - .48	.00 - .06

Appendix AM

Pooled Results of Multiple Regression Analyses for EARLI Measurement with Classroom Added as a Predictor

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.44	.15	2.90	<.01	.02 - .25	.02 - .25
Step 2						
Age	.04	.02	2.03	.04	.05 - .31	.01 - .12
Step 3						
Teacher 2	-.24	.25	-.96	.34		
Teacher 3	.23	.25	.91	.37		
Teacher 1	-.08	.31	-.27	.79	.07 - .34	.00 - .06

Appendix AN

Pooled Results of Multiple Regression Analyses for EARLI Pattern Recognition with Classroom Added as a Predictor

Predictors	*B*	*SE*	*t*	p-value	R^2	ΔR^2
Step 1						
Pretest	.59	.21	2.82	<.01	.02 - .43	.02 - .43
Step 2						
Age	.03	.02	1.34	.18	.06 - .43	.00 - .09
Step 3						
Teacher 2	-.01	.26	-.05	.96		
Teacher 3	-.14	.28	-.52	.61		
Teacher 1	.19	.32	.59	.55	.06 - .45	.00 - .10

Erin E. Reid

333 Marengo Road ◆ PA Furnace, PA 16865
(814) 689-2110 ◆ eer121@psu.edu

EDUCATION

Ph.D. **The Pennsylvania State University**, School Psychology, expected August 2010
M.Ed. **The Pennsylvania State University**, School Psychology, May 2007
B.S. **Mansfield University**, Psychology, December 2001

RESEARCH & WORK EXPERIENCE

Principal Investigator, *NumberFun Early Numeracy Skill Program*
Project Manager, *Early Arithmetic and Reading Learning Indicators (EARLI)*
Research Assistant, *Family Foundations: Coparenting and the Transition to Parenthood*
Instructor, *College of Education, The Pennsylvania State University*
Consultant, *The Second Mile*
School Psychologist Intern, *Tuscarora Intermediate Unit 11*
School Psychologist Intern, *Huntingdon Area School District*

PUBLICATIONS & PRESENTATIONS

DiPerna, J. C., Lei, P. W., Reid, E. E. (2007). Kindergarten predictors of mathematical growth in the primary grades: An investigation using the Early Childhood Longitudinal Study – Kindergarten Cohort. *The Journal of Educational Psychology*, 99, 369-379.

Reid, E. E., Morgan, P. L., DiPerna, J. C., & Lei, P. W. (2006). Development of measures to assess young children's early academic skills: Preliminary findings from a Head Start-university partnership. *Insights on Learning Disabilities, 3*(2), 25-38.

Reid, E. E., Miller, A., McGinnis, A. M., Goffreda, C., Reid, M. A., Myer, E. L., Hahn, Kathryn R., & Culler, E. (2009, February*). Construct validity of the Woodcock-Johnson III among adjudicated adolescents.* Poster session presented at the annual meeting of the National Association of School Psychologists, Boston, MA.

Reid, E. E., DiPerna, J. C., Morgan, P. L., Lei. P., & Wu, Q. (2008, November). *Early Arithmetic and Reading Learning Indicators: Assessing preschoolers' early academic skills using progress-monitoring measures.* Poster session presented at the annual meeting of the National Association for the Education of Young Children, Dallas, TX.

McGinnis A. M., DiPerna, J. C., Morgan, P. L., Lei, P., Reid, E. E., & Wu, Q. (2008, August). *An examination of sex differences in progress-monitoring performance in preschool students*. Poster session presented at the annual meeting of the American Psychological Association, Boston, MA.

Reid. E. E., & DiPerna, J. C. (2008, June). *Promoting early numeracy skill growth in Head Start children: NumberFun pilot study.* Poster session presented at the Head Start's Ninth National Research Conference, Washington, D.C.

DiPerna, J. C., Morgan, P. L., Lei., P. W., Reid, E. E., & Wu, Q. (2008, June). *The EARLI Literacy probes: Findings from year 3*. Poster symposium presented at the Head Start's Ninth National Research Conference, Washington, D.C.

DiPerna, J. C., Morgan, P. L., Lei., P. W., Reid, E. E., & Wu, Q. (2008, June). *The EARLI Literacy probes: Findings from year 3*. Poster symposium presented at the Head Start's Ninth National Research Conference, Washington, D.C.

Reid, E. E., & DiPerna, J. C. (2007, August*). Initial structural and concurrent validity evidence for the Teacher's Rating Scales of Early Academic Competence*. Poster session presented at the annual meeting of the American Psychological Association, San Francisco, CA.