Implementation of Early Numeracy Level 1 Intervention with First Grade Students

By

Melissa K. Trites

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Education Specialist
School Psychology

At

The University of Wisconsin-Eau Claire

May, 2020

Graduate Studies

The members of the Committee approve the thesis of

Melissa K. Trites presented on May 1st, 2020

____

Dr. Mary Beth Tusing, Chair

____

Dr. Melissa Coolong-Chaffin

____

Dr. Rosemary Battalio

APPROVED: ____
Dean of Graduate Studies

Implementation of Early Numeracy Level 1 Intervention with First Grade Students

By
Melissa K. Trites

The University of Wisconsin-Eau Claire, 2020
Under the Supervision of Dr. Mary Beth Tusing

This study evaluated the effectiveness of Early Numeracy Interventions Level 1 (Bryant et. al, 2011; ENI1) when used in a Montessori setting with first grade students. An ABAB design was used to evaluate the effectiveness of the intervention for the five participants. FastBridge measures of Number Identification, Place Value, and Decomposing were collected weekly. In addition, ENI1 Unit Checks were assessed before and after each unit to evaluate gains. Effect sizes were strong to poor in Number Identification, moderate to poor in Place Value and moderate to poor in Decomposing. Number Identification and Place Value showed the strongest results for most participants. Results indicated promise in using the ENI1 intervention within a nontraditional educational setting. Implications from the study are discussed in terms of ENI1's transportability including its use in a nontraditional educational setting, considerations for who implements the intervention, and whether the intervention is acceptable by teacher and participants in a Montessori setting.

Mary Beth Tusing, Ph.D., Thesis Advisor Date

To my parents, John and Judi Trites, who are my biggest encouragers and have always supported me in everything I do. Thank you for your love and support through the years.

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Mary Beth Tusing, for her support and guidance on this project. Thank you to the members of my thesis committee, Dr. Melissa Coolong-Chaffin and Dr. Rosemary Battalio, for their feedback and interest in this project. Lastly, thank you to the school that participated and allowed me to work with their students. I would not have been able to complete my thesis project without these individuals.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

The National Assessment of Educational Progress (NAEP) found that in 2015 only 40% of fourth-grade students, 33% of eighth grade students, and 25% of twelfth grade students in the nation performed at or above proficient levels in mathematics (NAEP, 2015). Early numeracy is a precursor to later math development. Therefore, when students struggle with early numeracy skills it is often the first sign of potential delays in math (Bryant, Bryant, Gersten, Scammacca, & Chavez, 2008). Given evidence that early intervention can prevent later learning difficulties (Jordan., Kaplan, Locuniak, & Ramineni, 2007; Bryant et al., 2008), more research on the effectiveness of early intervention programs for students with early numeracy delays are needed.

Early Numeracy

The intervention and assessment literature refer to early numeracy in a number of ways. Baglici and colleagues (2010) define early numeracy as understanding what numbers mean, demonstrating fluency and flexibility when using numbers, and having the ability to make quantity comparisons and perform mental calculations. Lembke and Foegen (2009) described early numeracy in two dimensions. The first involves basic number skills like counting, estimations, and number naming. The second involves the understanding of how numbers can be combined to make larger numbers or broken into smaller parts. Consistent with definitions of early numeracy, interventions targeting early math skills typically address skills related to counting, reading numbers, representing quantities as numbers, comparing and ordering numbers, understanding basic addition and subtraction, and using math with real world problems (Bryant et al., 2008; Conoyer, Foegen & Lembke, 2016; Mononen, Aunio, Koponen, & Aro, 2014).

The Common Core State Standards (Common Core State Standards Initiative, 2019) provide benchmarks for early numeracy skills that should be mastered in Kindergarten and first grade. By the end of Kindergarten, students should be able to count in sequence to 100 by ones and tens, understand quantities and number comparisons (e.g., greater than, less than, or equal to), understand addition within 10, understand subtraction within 5, and have beginning understanding of place value with numbers less than 20. By first grade, students should know how to add and subtract within 20; apply commutative, associative, identity, and distributive properties of addition and subtraction; count to 120 starting at any number; understand place value; and use place value and properties of operations to add and subtract.

Impact of Early Numeracy on Later Math Skills

Difficulties with early numeracy in early grades has long been associated with difficulties in mathematical knowledge in later childhood (Codding, Chan-Iannetta, Ferreira, & Volpe, 2011; Nelson & McMaster, 2019; Toll & Van Luit, 2012; Van Luit & Schopman, 2000). For example, Aunio, Heiskari, Van Luit, and Vuorio (2015) examined the development of early numeracy skills across a school year for Finnish Kindergarten students who were placed in a low-, average-, or high-performance group for early numeracy. Participants included 111 girls and 124 boys. Kindergarten early numeracy skills were assessed with the Finnish Early Numeracy Test (ENT; Aunio et al., 2006), which targets skills in comparing numbers, classifying numbers, one-to-one correspondence when counting, and accuracy when counting numbers. The researchers found that the low-performing group remained weak throughout the whole year. That is participants who entered Kindergarten behind in early numeracy skills showed growth but remained behind higher performing students at the end of the academic year.

Students who are better able to grasp early numeracy concepts including a basic understanding of numbers and addition and subtraction in Kindergarten are likely to have stronger math skills in second grade. Locuniak and Jordan (2008) examined Kindergarten predictors of math calculation fluency in second grade. Kindergarten assessments included measures of cognitive abilities, assessments of reading skills, and assessments of early math skills (i.e., counting, number naming, paper and pencil versus oral addition and subtraction problems, and story problems). Participants were from six different schools in the same school district. Early numeracy in Kindergarten was found to predict second grade calculation fluency better than other general predictors such as age, reading ability, oral vocabulary, or memory and spatial reasoning skills. Completing addition and subtraction calculations orally and having a better overall knowledge of numbers (i.e., counting, number naming skills) were the strongest predictors of second grade math calculation fluency.

Other researchers have reported similar findings. Kiss, Nelson and Christ (2019) completed a longitudinal study to evaluate the relationship between first grade early numeracy skills and third grade math performance on a state assessment of academic skills. Participants were from two school districts in the Midwest. First grade early numeracy skills were assessed with measures targeting skills in decomposing numbers (i.e., using two parts to make a whole), ordering numbers, reading numbers, identifying place value, answering orally presented addition and subtraction problems, solving story problems, and solving beginning addition and subtraction computations on paper. First grade students with significantly lower scores on early numeracy assessments did not meet proficiency standards on the state test in third grade. Further, decomposing, sequencing numbers, and basic addition and subtraction skills were the best predictors of third grade achievement.

Intervening with students who are having difficulties with early numeracy skills early on in their development will benefit their math achievement later on. Shanley, Clarke, Doabler, Kurtz-Nelson, and Fien (2017) examined whether remediating early numeracy skills for at-risk Kindergarteners supported their later math skill development. The researchers found that gains in early numeracy skills in Kindergarten were associated with higher mathematic achievement at the end of Kindergarten as well as during a first grade follow up. Therefore, early intervention is key in supporting later math development (Jordan et al., 2007; Bryant et al., 2008).

Supplemental Math Interventions to Support Early Numeracy

Several key educational organizations provide guidance on important components of math instruction. According to National Center on Intensive Intervention (National Center on Intensive Intervention at American Institutes for Research, 2019), supplemental math interventions should include explicit and systematic instruction, questions to expand mathematic knowledge, different models and ways of completing the problems, math vocabulary and symbol teaching, fluency building, and error analysis. According to What Works Clearinghouse (2013), there are five recommendations for teaching math to young children: 1) teaching numbers using a logical progression to build on skills; 2) teaching patterns, measurement, and data analysis using a logical progression to build on skills; 3) using progress monitoring to ensure skills are learned and retained; 4) teaching students to view and describe their surroundings mathematically; and 5) dedicating time each day to teaching math instruction. What Works Clearinghouse (2009) also emphasizes that interventions should focus on teaching whole number relationships, teaching skills explicitly, and using visual representations of math concepts. In addition, the development of fluency and automaticity with beginning math skills should be emphasized.

Recent meta-analyses have shown that the most effective early intervention programs typically include the elements described above. Mononen and colleagues (2014) completed a metanalysis of 19 math intervention studies that targeted early numeracy programs for four- to seven-year-olds at risk for math difficulties. The purpose was to review intervention effectiveness and summarize implementation variables like setting, duration of intervention, early numeracy content, professional support offered, and instructional design. The studies included in the review used a variety of instructional methods to improve early numeracy skills including explicit instruction with guided practice, peer-assisted tutoring, game and play based activities, and computer programs.

The most effective supplemental interventions in the review provided explicit instruction in targeted early numeracy skills, instruction in a small group or individual setting, and frequent monitoring of progress to determine if the intervention was effective in improving mathematics skills. Completing interventions in a small group or one on one setting produced higher effect sizes compared to whole class interventions. Duration of interventions varied and shorter interventions (i.e., less than twelve weeks) implemented within regular math instruction time or as a pullout session also significantly improved early numeracy skills. The intervention studies provided evidence that early numeracy interventions can effectively improve early numeracy skills in young students at risk for later mathematic difficulties.

Nelson and McMaster (2019) also completed a meta-analysis to evaluate the effects of early numeracy interventions for preschool and early elementary grades. Thirty-four studies were included. There were 52 treatment groups included in the meta-analysis. The studies included treatment groups that received interventions that supplemented core instruction (35%), treatment groups that received interventions that replaced part or all of the core math instruction, and treatment groups where the nature of the treatment condition was unclear (27%). The majority of interventions (69%) included two or more components of explicit instruction. The majority of the interventions focused on numbers and number relations skills, while fewer interventions focused on operation skills. Questions reviewed include: what is the mean effect of early numeracy interventions; how do treatment effects differ across participants, intervention, and outcome characteristics; which early numeracy domain was most investigated and produced the largest effect size; what were the frequency of skills in numeracy domains; and which variables accounted for the most between-studies variance for the total sample.

Results found larger treatment effects for interventions that were 8 weeks or shorter and smaller effect sizes for interventions that lasted longer than 8 weeks. The researchers stated that perhaps the smaller effects for interventions that lasted longer than 8 weeks can be explained by fadeout effects (i.e., strong effects initially, but those effects decreased as the intervention continued and the students in the conditions caught up to peers). Interventions with small group, flexible grouping, and one-to-one treatment groups showed moderate and large effects compared treatment groups using peer-assisted learning.

Early numeracy interventions included had strong treatment effects and the interventions were effective in improving early numeracy skills for participants. When considered together, the average effect across early numeracy interventions was moderate. Few studies provided detailed descriptions and examples of the lesson plans used for both treatment and control conditions. Few studies provided information regarding the implementation fidelity for different features of the intervention and the descriptions of how fidelity data was gathered. Larger effects occurred for interventions that included counting, comparing numbers, identifying numbers, and understanding number lines than those that did not include these concepts. Counting with one-to-one correspondence accounted for 51% of the variance in treatment effects, which shows the importance of including counting in early numeracy interventions. Early numeracy interventions were most effective for students with moderate risk for math delays than they were for students with significant math delays.

Interventions for Kindergarten and First Grade Students

Kindergarten Intervention Studies. The former meta-analysis included studies with preschool age children. Codding et al. (2011) evaluated a peer assisted, class-wide peer tutoring intervention, Kindergarten Peer Assisted Learning Strategies (Fuchs, Fuchs, & Karns, 2001), that targeted early numeracy skills for Kindergarten. Specific skills targeted were number naming skills, number comparison, and addition and subtraction. The intervention was delivered twice a week for 12 weeks. Participants were from six Kindergarten classrooms from two schools within the same school district. Most participants were English Language Learners and over half of the sample qualified for free and reduced lunch. Participants were randomly assigned to three groups: control, treatment, and treatment with goal setting and reinforcement. Dependent variables included the Test of Early Mathematics Ability (Ginsburg & Baroody, 2003), and the Tests of Early Numeracy from AIMSweb (NCS Pearson, Inc., 2020). The treatment group showed significant improvements over the control group; however, treatment with goal setting and reinforcement showed the greatest effects overall.

Toll and Van Luit (2012) evaluated the effectiveness of “The Road to Mathematics” (Van Luit & Toll, 2013), which is also an early numeracy intervention for Kindergarten students. The intervention included explicit and systematic instructional approaches targeting oral counting and skip counting, counting tangible objects, number naming, measuring, number lines, and simple calculations. Participants were in their second year of Kindergarten at six different schools in the Netherlands. Approximately 20% of the sample was from a low SES background. Participants were included in the intervention if they scored below the 50th percentile on the Early Numeracy Test – Revised (ENTR; Van Luit and Van de Rijt, 2009), which assessed counting skills and estimation skills. The program was delivered twice weekly for eight weeks in a small instruction group of three to four participants. Participants in the program had larger gains on the ENTR compared to students enrolled in the school but not participating in the program. In particular, Kindergarteners who scored below average on the pre-intervention assessment showed the largest gains after participation.

Finally, Dyson, Jordan, and Glutting (2011) also evaluated a number sense intervention for low-income Kindergarten students at risk for mathematic difficulties.

Participants were from an urban community, where 91% of students were on reduced priced lunch program. The intervention emphasized whole number concepts related to counting numbers, comparing numbers, and part-whole relationships with numbers. The program was delivered three times a week for 9 weeks. Lessons were scripted with explicit modeling of the skills presented. Participants completed the Number Sense Brief (Jordan et al., 2010) and the Applied Problems and Calculation subtests from the Woodcock Johnson III Test of Achievement. Participants in the treatment group had higher gains relative to those in a control group on the Number Sense Brief assessment and outperformed the control group on the WJ III subtests after the intervention.

Overall, interventions associated with gains in early numeracy for Kindergartners were provided in small groups and between 8-12 weeks in length. Peer assisted learning, modeling and incorporating practice, and explicit and systematic teaching methods were associated with gains in early numeracy. The studies with early numeracy gains addressed the skills of counting numbers, comparing numbers, and computing early addition and subtraction problems. Consistent with National Center on Intensive Intervention (National Center on Intensive Intervention at American Institutes for Research, 2019), the studies described used explicit and systematic instruction, math vocabulary, and fluency building.

First Grade Intervention Studies. Early Numeracy Intervention Level 1 (ENI1; Bryant, Bryant, Roberts, Vaughn, Pfannenstiel, Porterfield, and Gersten, 2011) was developed as a supplemental early numeracy intervention for first grade students. It focuses on the early numeracy skills of counting (i.e., counting principles and counting sequences), number knowledge and relationships (i.e., comparing and ordering numbers), grouping (i.e., composing and decomposing numbers), basic fact knowledge (i.e., fact families and related facts), and beginning story problems. The program incorporates explicit and systematic instructional techniques, along with modeling, thinking aloud, guided practice, and error correction.

Bryant, Bryant, Gersten, Scammacca, and Chavez (2008) evaluated the effectiveness of an earlier version of ENI that involved 64 fifteen-minute lessons on early numeracy concepts over a period of 18 weeks delivered by the researchers of the intervention. The lessons were completed with small groups of students and supplemented the students' core math curriculum. Lessons targeted the participants' ability to count numbers, to read numbers accurately, to compare the size of two numbers, to sequence or order numbers, and to add and subtract beginning math facts. Participants were English speaking, first grade students who scored at or below the 25^{th} percentile on the Texas Early Mathematics Inventory – Progress Monitoring (TEMI-PM). At post-test, participants did not make significant gains on the TEMI-PM. Researchers attributed poor outcomes to the short intervention periods. Researchers also speculated that students needed more practice opportunities with different representational ways (i.e., visual, abstract).

Bryant, Bryant, Roberts, Pfannenstiel, Porterfield, and Gersten (2011) evaluated a revised version of ENI, ENI1, for first grade students with math delays. The main intervention change was in delivery intensity. Lessons were expanded in length and the sequence of lessons was rearranged. Small groups of 3-5 first graders participated in the intervention 4 days per week for 25 minutes per session across 19 weeks. The teaching routines from the previous study of ENI1 were maintained. In this large-scale study, participants were from 10 different elementary schools in the same community. All participants were identified as at risk for math delays as determined by the TEMI-PM. Weekly progress monitoring assessments as well as the Stanford Achievement Test – Tenth Edition (SAT) and the Texas Early Mathematics Inventories – Outcome (TEMI-O) were used to evaluate outcomes. Participants were randomly assigned to control and treatment groups. At the end of the intervention period, participants in the treatment condition scored significantly higher on the TEMI-PM and TEMI-O compared to the control group. At the end of the school year, 45% of the treatment group students were no longer at significant risk on math screening assessments and 22% of the comparison students scored in the no risk range.

Bryant and colleagues (2011) suggested that future research on ENI1 should investigate whether general education teachers and interventionists are able to implement the intervention with fidelity. In each Unit of ENI1, the numbers used increase by a decade (i.e., Unit 1: 0-20, Unit 2: 0-30, Unit 3: 0-40, etc.). The researchers identified that future research needs to occur on students who have a slower response to the intervention and whether the decades increase too quickly for their understanding and learning.

Statement of the Problem

A number of early numeracy intervention programs, including ENI1, show promise in their ability to remediate early numeracy skills. Although there is a beginning evidence base for ENI1's effectiveness as a supplemental early numeracy intervention with first graders, more research is needed to support its use. In particular, the ENI1 program's transportability, that is, the program's effectiveness in a nontraditional educational setting, needs to be examined. Similarly, acceptability of using this program within a nontraditional context by the educators and the participants also needs to be examined. Lastly, expanding on transportability to determine if the program can be delivered with fidelity by trained interventionists.

Current Study

The purpose of this study was to extend research on ENI1 by evaluating its effectiveness when implemented in a Montessori environment. First grade Montessori students at-risk for math delays participated in the study. Interventionists were trained university students, rather than classroom teachers or study researchers. Interventionist dyads implemented the intervention three times per week. The following research questions were addressed:

1. Does ENI1 result in positive math outcomes for participants when implemented in the context of a Montessori educational setting? Specifically, do students show improvements in place value knowledge, number naming skills, and the ability to add/subtract parts from a whole. Additionally, do students' scores on ENI1 Unit Tests increase from Pre-Test to Post-Test?
2. What is the social acceptability of ENI1 when used as a supplemental program in a Montessori setting? Specifically, do teachers and participants rate ENI1 as acceptable for use in their Montessori school?

CHAPTER II

Method

Participants and Setting

Participants were five first grade females who attended a Montessori elementary school in the upper Midwest. The school enrolled approximately 289 students in pre-kindergarten through 5th grade and was part of a larger school district which served 11,367 students. Montessori students participated in multi-grade classrooms, which included Children's House (Pre-Kindergarten and Kindergarten), Elementary I (grades 1-3), and Elementary II (grades 4-5). The five participants in this study were from three different Elementary I classrooms. Demographic data from the school's website showed that 87% of students were White, 5% identified as representative of two or more races, 4% were Hispanic/Latino, 2% Black, 2% Asian, and less than 1% American Indian or Alaskan Native. Within the school, 8% of students were enrolled in special education and 24% qualified for free and reduced lunch. None of the participants received special education services.

Participants were selected for additional math intervention by the school's data team, which included teachers, school interventionists, and the building principal. The team considered both beginning of the school year academic screenings in math and teacher observations. The participants' teachers completed the fall administration of the AIMSweb Plus Tests of Early Numeracy (NCS Pearson, Inc., 2020). Participants also completed the Mathematics cluster of the Woodcock-Johnson IV Tests of Achievement (WJIV; Schrank, Mather, & Mcgrew, 2014) to evaluate their math skills prior to intervention. The Mathematics cluster includes two subtests. For Applied Problems, participants were to use math knowledge and math reasoning abilities to solve story problems that were read to them. For Calculations, participants completed a variety of math equations including addition and subtraction problems that were presented through paper and pencil. Trained school psychology graduate completed the WJIV assessments with participants.

Participants engaged in core math instruction in their Elementary I classroom. Math was taught using Montessori methods. Fidelity to core instruction was not monitored during this intervention. Concepts taught throughout the school year included numbers and operations in base ten including understanding of tens and ones, comparisons of number size, adding within 100, adding and subtracting multiples of 10, commutative and associative property of addition, math symbols (i.e., +, -, =), and determining the unknown whole number in an addition or subtraction equation. Math instruction also addressed algebraic thinking, word problems, geometry, and measurement and data.

Early Numeracy Intervention Level 1 (ENI1) sessions occurred three days a week and lasted from 45 to 60 minutes. One graduate student in school psychology lead each intervention session. A second graduate or undergraduate student was present during intervention to support student participation. The intervention sessions occurred in the afternoon at the end of the school day. Sessions were conducted in a small group at one table in the cafeteria of the elementary school. Another intervention group was in the cafeteria at the same time. Some behavioral management strategies (i.e., mystery motivator, token economy, prize box) were incorporated to keep the participants engaged and to support students' behavior during intervention time. These behavior strategies included positive reinforcement to promote appropriate behavior and to motivate participants to actively engage in the intervention. Two participants in particular had more significant behavioral challenges than the others.

Lola. Lola scored at the 6th percentile on the fall administration of the Test of Early Numeracy Composite from AIMSweb Plus. On the Mathematics cluster on the WJIV, Lola scored within the 2nd percentile. Throughout intervention, Lola was often quiet and benefitted from encouragement to participate. She had no behavioral problems during intervention time.

Megan. Megan scored at the 1st percentile on the fall administration of the Test of Early Numeracy Composite from AIMSweb Plus. On the Mathematics cluster on the WJIV, Megan scored within the 28th percentile. Throughout the intervention, Megan frequently appeared inattentive and benefitted from prompts to persist with tasks.

Nova. Nova scored at the 20th percentile on the fall administration of the Test of Early Numeracy Composite from AIMSweb Plus. On the Mathematics cluster on the WJIV, Nova scored within the 51st percentile. Throughout the intervention, Nova had no behavioral challenges. She was eager to participate and often was the first to answer questions during lessons. She typically completed independent work quickly and accurately.

Olive. Olive scored at the 1st percentile on the fall administration of the Test of Early Numeracy Composite from AIMSweb Plus. On the Mathematics cluster on the WJIV, Olive scored within the 10th percentile. Throughout the intervention, Olive had significant behavioral challenges including often putting her head down, looking at other groups in the room, and leaving intervention to sit at a different table. These behaviors became more frequent as the intervention continued throughout the school year. Behaviors improved somewhat when additional behavioral supports for the whole group were put into place.

Polly. Polly scored at the 18^{th} percentile on the fall administration of the Test of Early Numeracy Composite from AIMSweb Plus. On the Mathematics cluster on the WJIV, Polly scored within the 25^{th} percentile. Throughout the intervention, Polly also exhibited significant behavior challenges including refusal to complete work, pushing her papers away from her, pushing her pencil away from her, crying, turning her body away from the group, and putting her head down. Behavior challenges occurred almost every day. Some of the challenges improved when Polly was motivated by behavioral supports provided to the whole group.

Dependent Variables and Validity Measures

FastBridge Learning (FastBridge Learning, LLC, 2020) earlyMath progress monitoring tools and ENI1 Unit Tests were used to evaluate outcomes. FastBridge's Place Value, Number Identification, and Decomposing were administered weekly. These three assessments are suggested by FastBridge to use for progress monitoring with first grade students. They were also most similar to the early numeracy concepts taught during the intervention. The Unit Tests were given on the first day of the unit for the Pre-Test and again on the eighth day of the unit for the Post-Test. Therefore, the Unit Tests were given around every four to five weeks.

One progress monitoring tool was used during each session, resulting in each tool being administered once a week. Typically, Place Value was administered on Tuesdays, Number Identification on Thursdays, and Decomposing on Fridays. Some days were missed due to school vacations or cancellations. Assessments were completed by either graduate or undergraduate students trained in administration procedures.

Place Value. Place Value assessed the participants' understanding of place value for two- and three-digit numbers. The paper/pencil version of Place Value was group administered in the intervention setting. Participants had two minutes to complete as many problems on a worksheet as possible. There were two types of problems: 1) participants circled pictures of base 10 blocks that represented visually presented numbers on the worksheet, and 2) participants interpreted pictures of base 10 blocks and then printed the number represented by the blocks. Number sizes up to 120 were represented. Participants could earn up to 16 points per two minutes.

According to the FastBridge technical manual (Christ et al., 2018), Place Value's concurrent validity with the Measures of Academic Progress for Primary Grades (MAP) was $r = .82$. Fall Place Value scores were moderately correlated with spring performance on the Group Mathematics Assessment and Classification Evaluation (GMADE; ($r = .57$). Place Value has strong reliability (internal consistency $r = .90$; test retest $r = .77$; inter rater $r = .98$) and progress monitoring probes show strong alternative form reliability ($r = .82$). In the spring of first grade, students are considered at benchmark if they earn 4 or more points, which corresponds to the 40^{th} percentile of a national sample. According to 2019-20 FastBridge norms, average weekly rates of growth for first graders on Place Value were .06 points per week.

Number Identification. Number Identification assessed the participants' ability to read numbers correctly. The examiner administered the paper/pencil version of the assessment to the participants. Participants were presented with a sheet of paper with 96 numbers listed in no particular order. Numbers were as large as 119. The participants were to read as many numbers as they could within one minute. Participants could earn up to 96 points per minute.

Number Identification's concurrent validity with the MAP was $r = .50$. Fall Number Identification scores were moderately correlated with spring performance on the GMADE ($r = .59$). Number Identification has strong reliability (internal consistency $r = .90$; test retest $r = .91$; inter rater $r = .98$) and progress monitoring probes show strong alternate form reliability ($r = .92$). In the spring of first grade, students are considered at benchmark if they earn 42 or more points, which corresponds to the 40th percentile of a national sample. According to 2019-2020 FastBridge norms, average weekly rates of growth for first graders on Number Identification were .44 points per week (Christ et al., 2018).

Decomposing. Decomposing assessed the participants' beginning understanding of single digit addition and subtraction. The examiner administered the paper/pencil version of the assessment to the participants. The assessment was timed for one minute. Participants were presented with a sheet of paper with a numeral at the top of the page. Below the number, three additional numbers, either represented by dots or numerals, were presented. The examiner pointed to one of the three numbers and asked how many more to make the number at the top of the page. Participants were to use the whole number and one part of the number to identify the other part. Problems included numbers within 20. Participants could earn up to 24 points per minute.

Decomposing's concurrent validity with the MAP was $r = .56$. Fall Decomposing scores were moderately correlated with spring performance on the GMADE ($r = .59$).

Decomposing has strong reliability (internal consistency $r = .87$; test retest $r = .83$; inter rater $r = .93$) and progress monitoring probes show strong alternate form reliability ($r = .84$). In the spring of first grade, students are considered at benchmark if they earn 8 or more points, which corresponds to the 40th percentile of a national sample. According to 2019-2020 FastBridge norms, the average weekly rate of growth for first graders on Number Identification was .15 points per week (Christ et al., 2018).

ENI Pre- and Post-Unit Tests. The ENI1 Unit Tests were completed by participants before and after each unit. The Unit Tests assessed early numeracy skills including place value, number sequencing, identifying greater and less than, and simple addition and subtraction. The format of each Unit Test was the same. The test included five problems with addition and/or subtraction, five problems with place value, five problems with number sequencing, and five problems with more/less than. Unit Tests changed in difficulty by including larger numbers (e.g., the first Unit Test used mostly one- to two- digit numbers up to 40 while the second unit test used all two-digit numbers up to 50). Unit Tests were group administered and completed in the intervention setting. Interventionists read instructions to all participants to accommodate differences in reading abilities. Tests were not timed. Most participants took 15 minutes to complete the test. Participants were given one point for each correct answer, with a maximum of 20 points. With curriculum-based assessments, the desired level of performance is 95% correct.

Social validity. To assess perspectives regarding the effectiveness, appropriateness, and fairness of the intervention, participants and teachers completed an author-created social validity questionnaire at the completion of the study (see appendix

A). The teacher social validity scale was revised from Lane et al. (2009) and included 15-items. Respondents rated their level of agreement with each statement using a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The participant's scale included 7-items and asked respondents to rate their level of agreement with each statement using a Likert-type scale ranging from 1 (not true at all) to 4 (very true). Higher scores (with a possibility of 75 for the teacher scale and 28 for the participant scale) indicated higher social validity.

Procedural Integrity and Scoring Agreement. Fifty-four percent of the progress monitoring periods were video recorded and evaluated for fidelity of procedures and scoring agreement. The FastBridge Observing and Rating Administrator Accuracy (ORAA; FastBridge Learning, LLC, 2019) was used. Point-by-point interobserver agreement (IOA) and procedural integrity of the administration was calculated using the formula: [number of intervals of agreement / (number of intervals of agreement + number of intervals of disagreements)] X 100. Procedural integrity for administration of the progress monitoring probes ranged from 94% to 100%. The scoring agreement was 100% on the progress monitoring probes reviewed.

Treatment Integrity. All intervention sessions were video recorded. Procedural checklists were developed to evaluate procedural integrity for each intervention lesson and warm up activity (see appendix B). Checklists included steps the interventionist was to follow for each part (i.e., modeling, group practice, independent practice) of the lesson. Essential steps of each part of the lesson were listed. One intervention day per week was evaluated, with a different intervention day selected each week (e.g., Tuesday on one week, Thursday the next week, and then Friday the week after that). This cycle repeated throughout the intervention period. If there was a cancelled day of intervention, then the observer chose a different day that week to observe. In total, 57% of the intervention days were reviewed. Session integrity was calculated with the formula: [number of interval agreements / (total number of intervals)] X 100. Treatment integrity ranged from 72%-100% (M = 97%; Mdn = 100%) for each lesson and warm up activity.

Procedures

Baseline. Baseline lasted one week. During baseline (A1), behavioral expectations were taught, and participants played math games during intervention time. Participants completed all three FastBridge earlyMath measures three times during baseline. The WJIV was also completed at this time. Classroom instruction did occur during this time. The intervention was withdrawn (A2) for three weeks (two weeks of winter break and the first week participants returned to school after winter break). Classroom instruction did not occur during the two weeks for winter break but did occur once participants returned from break. Data was collected during the third week of the withdrawal phase.

Intervention. ENI1 covers early numeracy concepts including place value; number sequencing; identifying greater and less than; and simple addition and subtraction through part-part-whole activities, teaching math fact rules, and counting up and counting down strategies. Through the intervention period, six units were completed in total with participants. The units changed in difficulty by including progressively larger numbers (i.e., the first unit covers grouping within 10 and one- to two- digit numbers up to 40 while unit six uses one- to two- digit number up to 95). Each unit included 8 days of intervention. Unit 1 included 12 lessons and units 2-6 included 14 lessons. See Appendix C for the progression of the skills targeted in each unit and lesson. Units 1 and 2 included concepts of counting up and down, skip counting, and making groups of numbers. Units 3-6 included number sequencing, addition and subtraction, and place value. Materials (worksheets, flashcards, Fact Family Mats, Part-Part-Whole Mats, Relationships of 10 Mats) from the ENI1 intervention program (Bryant et al., 2011) were used. In addition, cubes and rods and connecting cubes were used throughout the intervention when indicated by the ENI1 lesson plans. The first intervention phase (B1) for all participants was implemented for five weeks. Units one and two were completed during this intervention period. The intervention was reinstated (B2) for fifteen weeks. Units three through six were completed during this intervention period.

School psychology graduate students and undergraduate psychology, communication sciences and disorders, and education majors served as interventionists. Each day of intervention, one graduate student and one to two undergraduate interventionists facilitated intervention. The graduate student served as a lead interventionist and the undergraduates supported participant participation by promoting on task behavior and providing corrective feedback. Graduate students were trained in the main ENI1 procedures at the beginning of the first intervention period. New interventionists were trained at the beginning of the second intervention period. Each week, all interventionists reviewed each lesson as a group to make sure they understood the lesson objectives, procedures, and error correction procedures. Weekly meetings also allowed for interventionists collaboration around participants' behavior and performance. Graduate students practiced the lessons before working with the participants.

Each intervention day, the lead interventionist welcomed the participants, reviewed behavioral expectations, and reviewed the day's schedule. Then, the ENI1 intervention materials were presented. Intervention days typically had a 5-minute warm up activity and then two short lessons lasting 20 to 30 minutes each. Warm up activities encouraged practice with previously practiced skills such as addition and subtraction with the use of flashcards and number sequencing with the use of number sequence cards. Interventionists followed the ENI1 lesson scripts when presenting instruction. Each ENI1 lesson first modeled the targeted skill with blocks or manipulatives, by writing problems on a white board, or using the intervention mats. During modeling, participants were to follow along by completing the steps of the activities with the interventionists (i.e., when the interventionists put a problem on the white board the participants also wrote the same problem on their white board). Then participants worked independently on a guided practice worksheet where interventionists monitored their work and provided corrective feedback and reteaching as needed. Last, participants completed an independent learning worksheet without assistance or feedback. After participants completed the independent practice, interventionists reviewed answers with each student.

Experimental Design and Analysis

To determine the effects of the ENI1 on participants' early numeracy skills, a single case ABAB withdrawal design was used for all participants. The design was selected to demonstrate a functional relation between the intervention and participants' performance on the outcome measures. This approach involved an initial baseline condition (A1), initial introduction of the ENI1 intervention package condition (B1), withdrawal of the intervention package condition (A2), and reimplantation of the intervention package condition (B2). Decisions about introduction and withdrawal of the

ENI1 intervention conditions were made based on the schedule of the school and the schedule of the intervention lessons.

Visual analysis of the participants graphs, and tables occurred to determine the level, variability, and trend of data. Effect sizes were also calculated through percentages of nonoverlapping data. Unit Checks were also analyzed through calculating percent correct for each Pre- and Post-Test.

CHAPTER III

Results

This study used a single case ABAB withdrawal design to evaluate the effects of Early Numeracy Intervention 1 (ENI1), a small group early numeracy intervention program, on the early numeracy skills of five first grade students at a Montessori elementary school. Specifically, ENI1 Unit Checks were examined for changes in percent correct from Pre-Test to Post-Test. Also, changes in participants' performance on FastBridge (FastBridge Learning, LLC, 2020) measures Number Identification, Place Value, and Decomposing after introduction of the intervention was evaluated. Visual analysis was used to examine changes in level, trend and variability in performance during intervention changes. Percentage of nonoverlapping data points (PND) was also evaluated. Finally, social acceptability ratings by teachers and participants was analyzed. Results for each participant are described below.

Table 1 summarizes each participant's percent correct on the Pre- and Post-Test Unit Checks by each Unit. During the first two Unit Checks, participants had lower percent corrects on the Pre-Test compared to the Post-Test. During Unit 1 and Unit 2, three participants showed increases from Pre-Test to Post-Test. For Units 5-6, three participants generally scored high on both Pre-Test and Post-Test. Ninety-five percent accuracy was seen on three participants on the Unit 3 Pre-Test, one out of four on the Unit 3 Post-Test, two out of four participants on the Unit 4 Pre-Test, three out of five on the Unit 4 Post-Test, three out of four participants on the Unit 5 Pre-Test, four out of four participants on the Unit 5 Post-Test, three out of five participants on the Unit 6 Pre-Test, and four out of five participants on the Unit 6 Post-Test. However, Megan's scores remained variable throughout the study and she consistently scored below 95% accuracy. If a participant was absent the day a Unit Test was given, the data was left blank.

Table 1

Percent Correct on Pre- and Post-Test Unit Checks

Units		Lola	Megan	Nova	Olive	Polly
Unit 1	Pre-Test	40%	40%		40%	50%
	Post-Test	55%	55%		75%	
Unit 2	Pre-Test	65%	60%		70%	50%
	Post-Test	90%		100%	80%	75%
Unit 3	Pre-Test	95%		95%	95%	80%
	Post-Test	90%	75%	100%		65%
Unit 4	Pre-Test	95%	85%	95%		55%
	Post-Test	95%	70%	95%	95%	60%
Unit 5	Pre-Test	95%	65%	100%	100%	
	Post-Test	95%		100%	100%	95%
Unit 6	Pre-Test	100%	60%	100%	90%	100%
	Post-Test	100%	55%	100%	95%	95%

Tables 2-4 summarize the participants' performance on Number Identification and Decomposing by providing the mean problems correct per minute (PCPM) and percent correct, and Place Value by providing the mean problems correct per two minutes (PCTM) and percent correct through each intervention phase (baseline, intervention, withdrawal, and reinstated intervention) for each participant.

Table 2

Number Identification Means and Ranges for Baseline and Intervention Phases

		A1	B1	A2	B2
Lola					
	Mean	24	33	36	51
	Range	(15-29)	(30-35)	(33-37)	(45-54)
	Mean % Correct	84%	86%	82%	90%
	Range	(68%-91%)	(83%-88%)	(79%-88%)	(85%-93%)
Megan					
	Mean	9	9	7	10
	Range	(8-10)	(8-9)	(6-7)	(7-14)
	Mean % Correct	60%	50%	40%	45%
	Range	(50%-67%)	(47%-53%)	(38%-41%)	(33%-70%)
Nova					
	Mean	30	33	40	50
	Range	(28-30)	(32-35)	(35-45)	(41-63)
	Mean % Correct	93%	92%	90%	93%
	Range	(93%-94%)	(91%-92%)	(86%-92%)	(87%-100%)
Olive					
	Mean	8	9	7	13
	Range	(7-9)	(8-9)	(6-9)	(8-18)
	Mean % Correct	30%	33%	40%	44%
	Range	(19%-38%)	(21%-44%)	(38%-41%)	(30%-56%)
Polly					
	Mean	15	25	26	28
	Range	(14-15)	(22-28)	(24-27)	(20-33)
	Mean % Correct	77%	92%	90%	93%
	Range	(76%-77%)	(89%-94%)	(82%-97%)	(78%-100%)

Table 3

Place Value Means and Ranges for Baseline and Intervention Phases

		A1	B1	A2	B2
Lola					
	Mean	0	2	6	9
	Range	(0-0)	(0-6)	(4-8)	(6-13)
	Mean % Correct	0%	17%	43%	72%
	Range	(0%-0%)	(0%-60%)	(27%-67%)	(43%-93%)
Megan					
	Mean	0	0.2	0.3	2
	Range	(0-0)	(0-1)	(0-1)	(0-5)
	Mean % Correct	0%	2%	7%	18%
	Range	(0%-0%)	(0%-8%)	(0%-20%)	(0%-50%)
Nova					
	Mean	3	7	9	12
	Range	(3-3)	(5-9)	(7-11)	(7-15)
	Mean % Correct	51%	87%	90%	86%
	Range	(43%-60%)	(50%-100%)	(88%-92%)	(70%-100%)
Olive					
	Mean	0	4	5	10
	Range	(0-0)	(0-8)	(4-6)	(4-13)
	Mean % Correct	0%	33%	31%	84%
	Range	(0%-0%)	(0%-50%)	(25%-38%)	(50%-100%)
Polly					
	Mean	0	4	5	6
	Range	(0-0)	(0-7)	(3-7)	(4-9)
	Mean % Correct	0%	34%	80%	60%
	Range	(0%-0%)	(0%-88%)	(60%-100%)	(8%-88%)

Table 4

Decomposing Means and Ranges for Baseline and Intervention Phases

		A1	B1	A2	B2
Lola					
	Mean	5	5	2	5
	Range	(3-7)	(2-7)	(1-3)	(1-7)
	Mean % Correct	67%	41%	22%	45%
	Range	(50%-80%)	(22%-58%)	(17%-33%)	(17%-100%)
Megan					
	Mean	1	2	3	4
	Range	(0-2)	(0-3)	(3-4)	(2-7)
	Mean % Correct	23%	22%	37%	37%
	Range	(0%-50%)	(0%-33%)	(33%-44%)	(22%-50%)
Nova					
	Mean	6	7	8	12
	Range	(5-6)	(5-9)	(6-9)	(8-15)
	Mean % Correct	62%	58%	68%	76%
	Range	(55%-75%)	(50%-69%)	(50%-81%)	(61%-100%)
Olive					
	Mean	5	6	3	5
	Range	(3-6)	(3-8)	(2-4)	(2-10)
	Mean % Correct	40%	52%	33%	42%
	Range	(30%-46%)	(50%-53%)	(22%-44%)	(22%-58%)
Polly					
	Mean	3	3	3	4
	Range	(2-3)	(2-4)	(2-5)	(2-7)
	Mean % Correct	24%	30%	32%	40%
	Range	(20%-27%)	(22%-44%)	(22%-42%)	(22%-58%)

Table 5 summarizes the percentage of nonoverlapping data (PND) during both intervention periods for all participants for each FastBridge outcome measure.

Table 5

Percentage of Nonoverlapping Data (PND) from Baseline to Intervention

Participant	Number Identification	Place Value	Decomposing
Lola	80%	40%	0%
Megan	30%	29%	42%
Nova	33%	50%	71%
Olive	67%	69%	17%
Polly	50%	15%	21%

Lola. Performance on the three curriculum-based measures for Lola are shown in Figure 1. Baseline (A1) Number Identification (NID) showed a slight upward trend and ranged from 15 to 29; however, the first intervention phase (B1) NID trend was positive relative to the last two baseline data points (range: 30-35). Lola's mean NID increased from 24 (A1) to 33 (B1). During the withdrawal phase (A2), Lola's mean NID of 36 was similar to B1. The reintroduction of the intervention (B2) resulted in a higher mean NID of 51. No changes occurred in the trend from B1 to A2. In B2, Lola's trend flattened, which may have been due to a ceiling effect on the dependent measure. By the end of the intervention period, Lola's NID scores were in the benchmark range. Variability in B1 (range: 30-35) and A2 (range: 33-37) were similar. Then, variability increased in B2 (range: 45-54), which reflects Lola's NID growth over the longer intervention period. Lola showed moderate levels of accuracy during all intervention phases (M = 86% and

90%) and baseline phases (M = 84% and 82%) on NID probes. Effect size was strong for NID (PND = 80%).

A1 Place Value (PV) scores showed a stable trend and ranged from 0 to 0 while the B1 PV trend was positive (range: 0-6). Lola's mean PV increased from 0 (A1) to 2 (B1). During A2, Lola's mean (M = 6) PV increased compared to B1. The reintroduction of B2 resulted in a higher mean of PV of 9. A change in trend occurred from B1 to A2 and again when the intervention was reinstated (B2). PV showed a slight positive trend in B2. By the end of the intervention period, Lola's PV scores were in the benchmark range. No variability occurred during A1 (range: 0-0); however, a change in variability occurred during B1 (range: 0-6). Variability in B1 and A2 (range: 4-8) were similar. Then, variability slightly increased in B2 (range: 6-13), which reflects Lola's PV growth over the longer intervention period. Lola showed poor levels of accuracy during intervention phases (M = 71% and 72%) and very poor levels during baseline phases (M = 0% and 43%) on PV probes. Effect size was moderate for PV (PND = 40%).

A1 Decomposing (DC) showed a slight upward trend and ranged from 3 to 7 and the B1 DC trend was positive as well (range: 2-7). Lola's mean DC were the same from A1 (M = 5) to B1 (M = 5). During A2, Lola's mean decreased (M = 2) compared to B1. The reintroduction of B2 resulted in a higher DC mean of 5. A decrease in trend occurred from B1 to A2. Lola's trend was flat during B2. By the end of the intervention period, Lola's DC scores were not at benchmark. Variability in B1 (range: 2-7) and A2 (range: 1-3) were similar and proceeded by similar variability in B2 (range: 1-7), which suggests limited change in DC skill as a result of the intervention. Lola showed poor levels of accuracy during all intervention phases (M = 41% and 45%) and baseline phases (M = 67% and 22%) on DC probes. Effect size was poor for DC (PND = 0%).

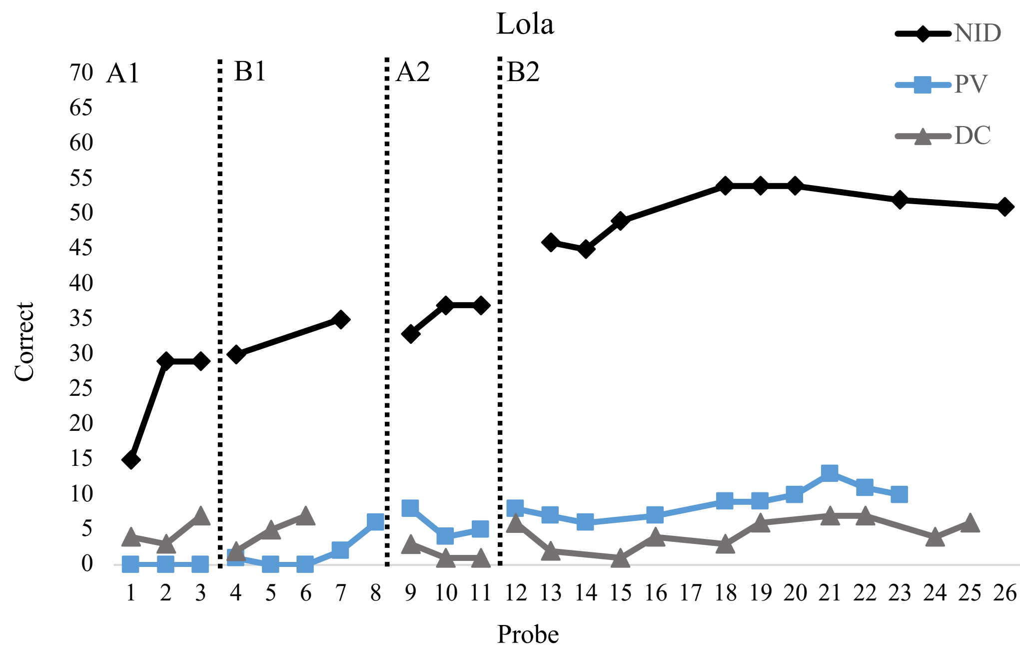


Figure 1. Problems correct per minute for Number Identification (NID) and Decomposing (DC), and problems correct per two minutes for Place Value (PV) for Lola.

Megan. Performance on the three curriculum-based measures for Megan are shown in Figure 2. Baseline (A1) Number Identification (NID) showed a slight downward trend and ranged from 8 to 10. The first intervention phase (B1) trend remained flat (range: 8-9) compared to the A1 phase. Megan's mean NID stayed the same from A1 (M = 9) to B1 (M = 9). During the withdrawal phase (A2), Megan's mean NID of 7 decreased compared to B1. The reintroduction of the intervention (B2) resulted in slightly higher mean NID of 10. No changes occurred in the trend from B1 to A2. In B2, Megan's trend had a slow, slight upward trend. By the end of the intervention period, Megan's NID scores were not in the benchmark range and fell in the high-risk category. Variability in B1 (range: 8-9) and A2 (range: 6-7) were similar. Then, variability increased in B2 (range: 7-14), which reflects Megan's slight, slow growth over the longer intervention period. Megan showed poor levels of accuracy during all intervention phases (M = 50% and 45%) and baseline phases (M = 60% and 40%) on NID probes. Effect size was poor for NID (PND = 30%).

A1 Place Value (PV) showed a stable trend and ranged from 0 to 0, and the B1 PV trend was stable as well (range: 0-1). Megan's mean PV stayed the same from 0 (A1) to 0.2 (B1). During A2, Megan's mean (M = 0.3) PV stayed stable compared to B1 (M = 0.2). The reintroduction of B2 resulted in a slightly higher mean of PV of 2. The trend between B1 and A2 stayed the same; however, there was a slight change in trend when the intervention was reinstated (B2). Megan's PV data showed a slight positive trend in B2. By the end of the intervention period, Megan's PV scores were only in the benchmark range twice. No variability occurred during A1 (range: 0-0), B1 (range: 0-1), nor A2 (range: 0-1). A slight change in variability occurred during B2 (range: 0-5) compared to A2, which reflects some PV growth over the longer intervention period. Megan showed poor levels of accuracy during all intervention phases (M = 2% and 18%) and baseline phases (M = 0% and 7%) on PV probes. Effect size was poor for PV (PND = 29%).

A1 Decomposing (DC) showed a downward trend and ranged from 0 to 2; however, the B1 DC trend was flat (range: 0-3). Megan's mean DC were similar from A1 (M = 1) to B1 (M = 2). During A2, Megan's mean stayed similar (M = 3) compared to B1, and the reintroduction of B2 resulted in a similar DC mean of 4. The trend remained flat from B1 to A2. Megan's trend slightly increased during B2, but then started to decrease again towards the end of the phase. By the end of the intervention period, Megan's DC scores were not at benchmark. Variability remained similar in all four phases: A1 (range: 0-2), B1 (range: 0-3), A2 (range: 3-4), and B2 (2-7), which suggests little changes in DC skills as a result of the intervention. Megan showed poor levels of accuracy during all intervention phases (M = 22% and 37%) and baseline phases (M = 23% and 37%) on DC probes. Effect size was moderate for DC (PND = 42%).

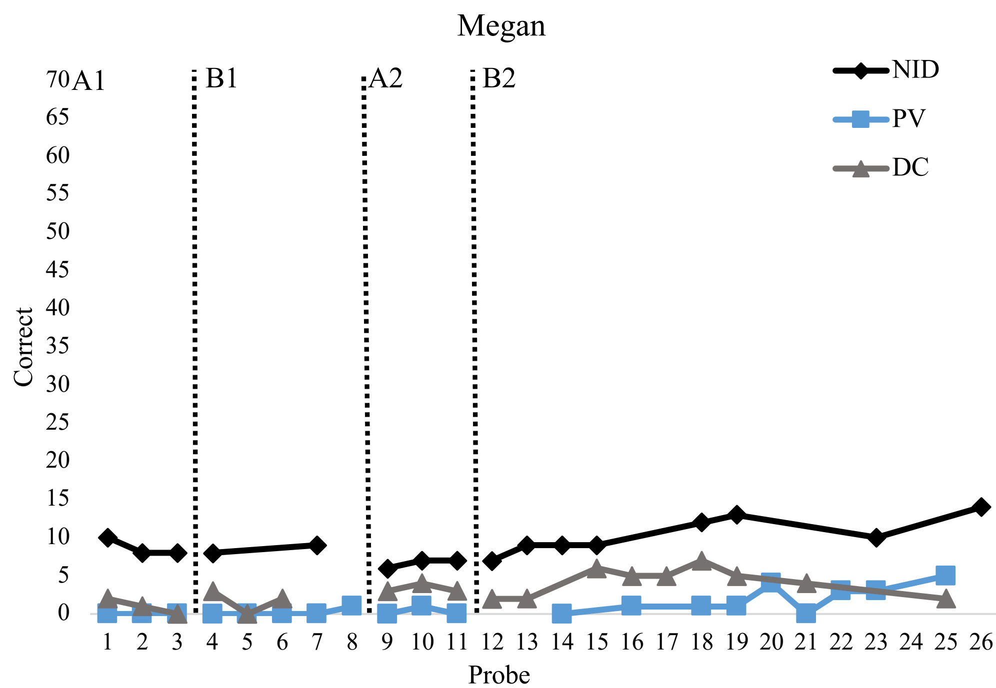


Figure 2. Problems correct per minute for Number Identification (NID) and Decomposing (DC), and problems correct per two minutes for Place Value (PV) for Megan.

Nova. Performance on the three curriculum-based measures for Nova are shown in Figure 3. Baseline (A1) Number Identification (NID) showed a slight upward trend and ranged from 28 to 30, and the first intervention phase (B1) NID trend was positive as well (range: 32-35). Nova's mean NID had a slight increase from 30 (A1) to 33 (B1). During the withdrawal phase (A2), Nova's mean NID increased (M = 40) compared to B1. The reintroduction of the intervention (B2) resulted in a higher mean of NID of 50. A change occurred in trend from increasing (B1) to decreasing (A2). In B2, Nova's trend increased compared to A2. By the end of the intervention period, Nova's NID scores were in the benchmark range. A change in variability occurred between B1 (range: 32-35) and A2 (range: 35-45). Then, variability increased in B2 (range: 41-63), which reflects Nova's NID growth over the longer intervention period. Nova showed moderate levels of accuracy during all intervention phases (M = 92% and 93%) and baseline phases (M = 93% and 90%) on NID probes. Effect size was poor for NID (PND = 33%).

A1 Place Value (PV) scores showed a stable trend and ranged from 3 to 3; however, the B1 PV trend (range: 5-9) was positive compared to the A1. Nova's mean PV increased from 3 (A1) to 7 (B1). During A2, Nova's mean (M = 9) stayed the same compared to B1. The reintroduction of B2 also resulted in a higher mean of PV of 12. A change in trend occurred from B1 to A2 and again when the intervention was reinstated (B2). PV showed a positive trend in B2 and decreasing trend in A2. By the end of the intervention period, Nova's PV scores were in the benchmark range. No variability occurred during A1 (range: 3-3); however, a change in variability occurred during B1 (range: 5-9). Variability increased in B1 to A2 (range: 7-11); then, variability increased again during B2 (range: 7-15), which reflects Nova's PV growth over the study phases. Nova showed moderate levels of accuracy during all intervention phases (M =87% and 86%) and baseline phases (M = 51% and 90%) on PV probes. Effect size was moderate for PV (PND = 50%).

A1 Decomposing (DC) showed stable trend and ranged from 5 to 6; however, the first data point in B1 was higher followed by a decrease in trend (range: 5-9). Nova's mean DC had a slight increase from A1 (M = 6) to B1 (M = 7). During A2, Nova's mean also slightly increased (M = 8) compared to B1. The reintroduction of B2 resulted in a higher DC mean of 12. A decrease in trend occurred from B1 to A2; however, Nova's trend slowly increased during B2. By the end of the intervention period, Nova's DC scores were at benchmark range. Variability in B1 (range 5-9), A2 (range: 5-9), and B2 (range: 6-9) were similar. Variability increased in B2 (range: 8-15), which suggests Nova's DC growth over the longer intervention period. Nova showed poor levels of accuracy during all intervention phases (M = 58% and 76%) and baseline phases (M = 62% and 68%) on DC probes. Effect size was moderate for DC (PND = 71%).

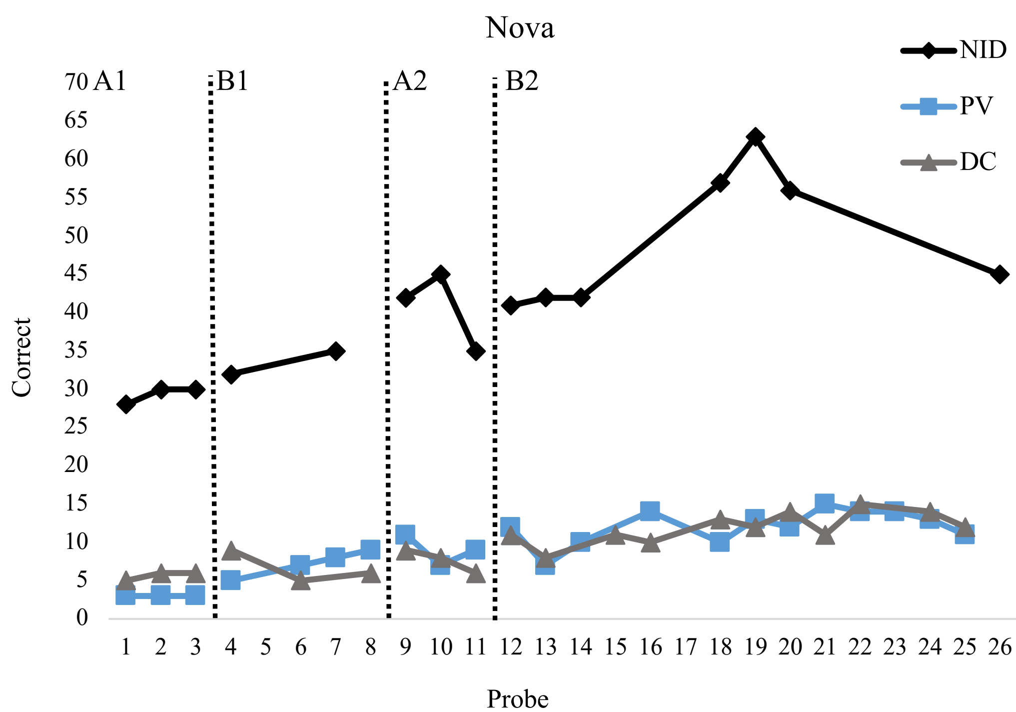


Figure 3. Problems correct per minute for Number Identification (NID) and Decomposing (DC), and problems correct per two minutes for Place Value (PV) for Nova.

Olive. Performance on the three curriculum-based measures for Olive are shown in Figure 4. Baseline (A1) Number Identification (NID) showed a slight downward trend and ranged from 7 to 9. In the first intervention phase (B1) NID's trend remained flat (range: 8-9). Olive's mean NID stayed the same, from 8 (A1) to 9 (B1). During the withdrawal phase (A2), Olive's mean NID of 7 was a decrease from B1. The reintroduction of the intervention (B2) resulted in a higher mean NID of 13. Olive's data starts trending upward in A2 and continues to increase in B2. By the end of the intervention period, Olive's NID scores were not in the benchmark range. Variability in B1 (range: 8-9) and A2 (range: 6-9) were similar. Then, variability increased in B2 (range: 8-18), which reflects Olive's NID growth over the longer intervention period. Olive showed very poor level of accuracy during all intervention phases (M = 33% and 44%) and baseline phases (M = 30% and 40%) on NID probes. Effect size was moderate for NID (PND = 67%).

A1 Place Value (PV) scores were flat and ranged from 0 to 0; however, the first intervention phase (B1) showed a positive trend (range: 0-8). Olive's mean PV increased from 0 (A1) to 4 (B1). During A2, Olive's mean (M = 5) stayed the same compared to B1, and the reintroduction of B2 resulted in a higher mean of PV of 10. A change in trend occurred from B1 to A2 and again when the intervention was reinstated (B2). PV showed a slight positive trend in B2. By the end of the intervention period, Olive's PV scored were in the benchmark range. No variability occurred during A1 (range: 0-0); however, a change in variability occurred during B1 (range: 0-8). Variability in B1 and A2 (range: 4-6) were similar. Then, variability increased in B2 (range: 4-13), which reflects Olive's growth over the longer intervention period. Olive showed poor accuracy during B1 (M = 33%) and then improved to moderate range in B2 (M = 84%). Olive showed poor levels of accuracy during both baseline phases (M = 0% and 31%). Effect size was moderate for PV (PND = 69%).

A1 Decomposing (DC) showed a slight upward trend and ranged from 3 to 6, and the B1 trend was positive as well (range: 3-8). Olive's mean DC slightly stayed the same from 5 (A1) to 6 (B1). During A2, Olive's mean decreased (M = 3) compared to B1 and her overall trend also changed directions. The reintroduction of B2 resulted in higher DC mean of 5 compared to A2. A slight upward trend occurred during B2. By the end of the intervention period, Olive's DC scores were at the benchmark range only once. Variability in A2 (range: 2-4) slightly decreased compared to B1 (range: 3-8). Then, variability slightly increased in B2 (range: 2-10) compared to A2, which reflect some but limited changes in DC skills as a result of the intervention. Olive showed poor levels of accuracy during all intervention phases (M = 52% and 42%) and baseline phases (M = 40% and 33%) on DC probes. Effect size was poor for DC (PND = 17%).

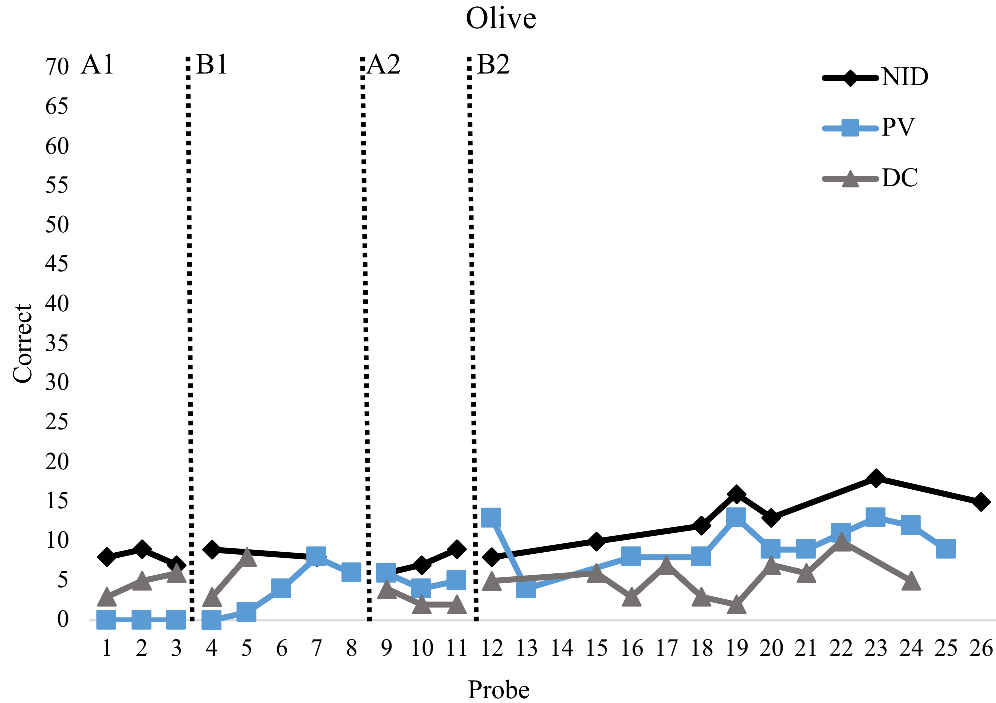


Figure 4. Problems correct per minute for Number Identification (NID) and Decomposing (DC), and problems correct per two minutes for Place Value (PV) for Olive.

Polly. Performance on the three curriculum-based measures for Polly are shown in Figure 5. Baseline (A1) Number Identification (NID) showed a stable trend and ranged from 14 to 15. Polly's performance showed a positive trend upward in the first intervention phase (B1) NID trend (range: 22-28) compared to A1. Her mean NID increased from 15 (A1) to 25 (B1). During the withdrawal phase (A2), Polly's mean NID of 28 was a slight increase compared to B1. The reintroduction of the intervention (B2) resulted in a slightly higher mean NID of 28. No changes occurred in the trend from B1 to A2. In B2, Polly's trend flattened throughout the phase. By the end of the intervention period, Polly's NID scores were not in the benchmark range. Variability in B1 (range: 22-28) and A2 (range: 24-27) were similar. Variability slightly increased in B2 (range: 20-

33), which reflects some NID growth over the longer intervention period. Polly's last data point during the B2 phase was lower. Without that data point, interpretations may be different. Polly showed moderate levels of accuracy during both intervention phases (M = 92% and 93%). Polly showed poor levels of accuracy during A1 (M = 77%), but then increased to moderate levels of accuracy during A2 (M = 90%). Effect size was moderate for NID (PND = 50%).

A1 Place Value (PV) scores remained flat and ranged from 0 to 0; however, the B1 PV trend was positive (range: 0-7). Polly's mean PV increased from 0 (A1) to 4 (B1). During A2, Polly's mean (M = 5) PV stayed the same compared to B1. The mean during the reintroduction of B2 (M = 6) also stayed the same compared to A2. A slight decrease in trend occurred in A2 compared to B1 and then a flat trend occurred when the intervention was reinstated (B2). By the end of the intervention period, Polly's PV scores were in the benchmark range twice. No variability occurred during A1 (range: 0-0); however, a change in variability occurred during B1 (range: 0-7). Variability in B1 and A2 (range: 3-7) were similar. Then, variability slightly increased in B2 (range: 4-9), which reflects Polly's slight PV growth over the intervention period. Polly showed poor levels of accuracy during all intervention phases (M =34% and 60%) and baseline phases (M = 0% and 80%) on PV probes. Effect size was poor for PV (PND = 15%).

A1 Decomposing (DC) showed a stable trend and ranged from 2 to 3 and the B1 DC trend remained flat (range: 2-4). Polly's mean DC were similar from A1 (M = 3) to B1 (M = 3) to A2 (M = 3). The reintroduction of B2 also resulted in a similar DC mean of 4. A similar trend occurred from B1 to A2. Polly's trend was also flat during B2, which shows no growth across the study. By the end of the intervention period, Polly's

DC scores were not at benchmark. Variability in B1 (range: 2-4), A2 (range: 2-5), and B2 (range: 2-7) were similar, which suggests limited changes in DC skills as a result of the intervention. Polly showed very poor levels of accuracy during all intervention phases (M = 30% and 40%) and baseline phases (M = 24% and 32%) on DC probes. Effect size was poor for DC (PND = 21%).

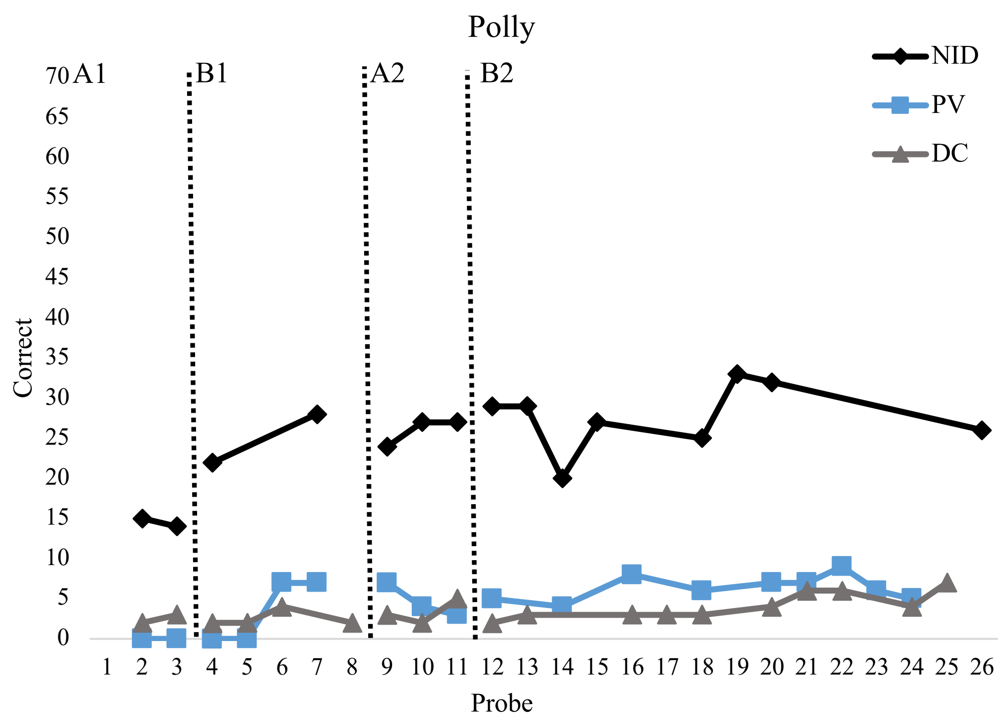


Figure 5. Problems correct per minute for Number Identification (NID) and Decomposing (DC), and problems correct per two minutes for Place Value (PV) for Polly.

All participants completed the social acceptability measure. The results of the participant social validity questionnaire were very favorable. For the first statement ("The math practice I did this year was good for me"), the mean score was 3.8 (SD = 0.45); for the second statement ("I liked the way I practiced math this year"), it was 3.8 (SD = 0.45); for the third ("I think my math skills improved this year"), it was 3.8 (SD = 0.45);

for the fourth (“I think other kids would think this math program is helpful”), it was 3 (SD = 1.41); for the fifth (“I would want to work on math with university students again”), it was 3.2 (SD = 1.30); for the sixth (“This program helped me understand math better”), it was 4 (SD = 0.0); and for the last statement (“Overall, I think the math work I did here was helpful”), it was 3.2 (SD = 1.30).

Only one of three teachers completed the teacher social acceptability measure. Table 6 shows the responses provided by the teacher. Overall, the results of the social validity questionnaire filled out by the teacher were also favorable.

Table 6

Scores of the Teacher Acceptability Survey

Question	Score
1. This is an acceptable intervention for students at a Montessori elementary school.	Strongly Agree
2. Most Montessori teachers would find this intervention appropriate.	Agree
3. I would suggest the use of this intervention to other Montessori teachers.	Strongly Agree
4. The intervention is appropriate to meet a Montessori school’s need and mission.	Agree
5. Most Montessori teachers would find this intervention suitable for the described purpose and mission.	Agree
6. This intervention would not result in negative side-effects for the students.	Strongly Agree
7. This intervention is consistent with those I have used or are offered in this school setting.	Strongly Agree
8. The intervention is a fair way to fulfill the intervention purposes.	Strongly Agree
9. This intervention is reasonable to meet the stated purposes.	Strongly Agree
10. Overall this intervention would be beneficial for a variety of Montessori elementary school students.	Strongly Agree

CHAPTER IV

Discussion

This study examined the effectiveness of a supplemental intervention program, Early Numeracy Intervention Level 1 (ENI1; Bryant et al., 2011), when implemented in a Montessori setting. The first objective was to determine if ENI1 would result in positive math outcomes when implemented in a nontraditional school context. The second objective was to determine the social acceptability of ENI1 being used within a Montessori context with teachers and participants.

Effectiveness of Early Numeracy Interventions Level 1

Previous research has found that ENI1 is effective for increasing early numeracy skills when comparing a treatment condition to a control condition (Bryant et al, 2008; Bryant et al., 2011). ENI1 intervention components and targets were consistent with features found in previous meta-analyses (Mononen and colleagues, 2014; Nelson & McMaster, 2019) that also increased early numeracy skills in students. In other studies of early numeracy interventions, larger effects have been found for interventions that incorporate number sequencing, place value, counting up and down, comparing numbers, and identifying number (Nelson & McMaster, 2019), which are also features of ENI1. Therefore, ENI1 includes intervention components that have previously increased early numeracy skills in students.

In this study, an ABAB single case design was used to determine whether ENI1 was effective in increasing early numeracy skills with participants from a nontraditional school. Overall, many participants made gains across most dependent variables in the study. Specifically, four participants made gains in Number Identification, four participants made gains in Place Value, and one participant made gains in Decomposing. More gains were seen during the Intervention 2 phase compared to the Intervention 1 phase. Throughout the study, four out of five participants overall scored well on the Unit Checks. One participant's scores on the Unit Checks were variable throughout the study.

Number Identification. The ENI1 intervention appeared to have a positive impact on Number Identification (NID) skills for most participants. Four of the five participants made gains in number reading fluency from the beginning of the year to the end of the year. In the ABAB design, noticeable differences in the level and trend of performance was evident for four out of five participants during intervention phases. Accuracy of scores ranged from poor to moderate levels. As the participants were in intervention, four participants' accuracy with NID increased. Effect sizes were moderate to strong for all participants.

Place Value. Place Value (PV) appeared to be the strongest area of growth when all data was considered. Four out of five participants made gains in their mean scores from the beginning of the year to the end of the year. Four out of five participants scored at the benchmark range at the end of the year. Three out of five participants' trends increased during the intervention phases. Two out of five participants' levels changed throughout the study. Accuracy of scores ranged from poor to moderate levels. All participants' accuracy scores increased as they were in the intervention. Effect sizes were poor to moderate.

Decomposing. Participants showed minimal gains in decomposing skills as assessed by the Decomposing measure from FastBridge. Only one participant showed noticeable gains in mean scores, changes in level, and increases in trend from the beginning to the end of the study. Results for only two participants showed moderate effect sizes, effect sizes for the remaining participants were poor and ranged from 0% to 21%.

Pre- and Post-Tests. Two trends were evident in participant performance on the Unit Tests. One is that most students showed improvements from pre- to post-test at the beginning of the study. Second, by the end of the intervention, most students were scoring well even on the pre-test. Potentially, the repetitive nature of the Unit Checks and the lessons in ENI1 may have impacted the participants' performance. There was no statistical control, however, so we cannot assume that the growth in these tests were because of the ENI1 program alone. Other variables that could have impacted the growth include classroom instruction and the participants becoming familiar with the tests and lessons.

Transportability and Acceptability in a Nontraditional Educational Setting

Fidelity of intervention administration was collected. The results showed that ENI1 can be implemented with fidelity by individuals with a variety of backgrounds and individuals besides the researchers. Transportability of an intervention means implementing the intervention under conditions of the local level or with community practitioners (Schoenwald & Hoagwood, 2001). Rather than the researchers implementing ENI1 in the Bryant et al. (2011) study, this study had graduate students implement ENI1. The high fidelity of intervention administration in this study, therefore, provides evidence for the transportability of ENI1 to be implemented by trained staff in nontraditional educational settings.

Overall, the participants and the teacher who filled out the social validity measures responded that ENI1 was beneficial and acceptable for use in a Montessori setting. Acceptability is important for interventions so teachers will implement the intervention and students will participate in the intervention. Treatment acceptability is also important because it helps researchers understand whether others found the intervention to be successful.

A benchmark for a measure is a standard that is set for that skill set. In this study, two participants were in the benchmark range for Number Identification at the end of the study. Four out of five participants were consistently in the benchmark range for Place Value at the end of the study. One participant was in the benchmark range for Decomposing at the end of the study. Nelson and McMaster (2019) and Kiss, Nelson, and Christ (2019) found that better performance on Number Identification in early grades was found to be important for success in later math development. Kiss, Nelson, and Christ (2019) also found that the skills needed to be successful with Place Value (i.e., understanding base 10 blocks visually and number representation) and Decomposing (i.e., using two parts to make a whole) were needed to be successful for later math development. It is promising that many of the participants scored in the benchmark range at the end of the study, because this level of performance is associated with less risk for future math challenges.

Implications for Research

ENI1 has many of the intervention components common in effective early numeracy programs (Mononen et al., 2014). Therefore, it is not surprising that ENI1 was associated with participant gains in identifying numbers and in place value skills.

However, the intervention did not appear to be effective for all students or for all early numeracy outcomes studied. Further, causal interpretations of the outcome data as indicators of participant gains in early numeracy skills due to the ENI1 program are more difficult to determine. There are several possible explanations for the current outcomes that have implications for future research.

Experimental Design. An ABAB design includes a baseline phase followed by an intervention phase, then a withdrawal phase occurs where the intervention is removed. Lastly, the intervention is reintroduced for the Intervention 2 phase. A strength of an ABAB design for academic data includes incorporating a direct replication component that allows for stronger experimental control (Riley-Tillman & Burns, 2009). Gains were seen by many participants in this study. The causality of those gains, however, are harder to determine. An ABAB design is meant to control for outside influences on the outcome being evaluated, because there is the assumption that the only thing that changes during the A phases is the presence of the intervention. This is accounted for in this study for the first AB phase, but then the replication of the effect is confounded by the fact that part of the second A phase was the holiday break. Not only was the intervention removed during this A phase, but during holiday break, classroom instruction was also removed during this phase. Therefore, a direct replication of the first AB phase did not occur.

A related concern for the ABAB design of this study is decision-making regarding baseline data. Another strength of the ABAB design is that data collection can be used to make decisions about when the intervention should be introduced and removed to test for an effect. If the intervention is associated with changes in performance relative to baseline or withdrawal phases, it is defensible to say that the intervention was the cause for the change (Riley-Tillman & Burns, 2009). Visual analysis of the data to decide when to move from a stable baseline to intervention did not occur in this study. Instead, due to time restraints and the need to start intervention as soon as possible, one week of baseline which included only three data points occurred.

Intervention Intensity. In addition to limitations in fidelity to the experimental design of the study that negatively impacted the ability to make cause interpretations of the data, the current study may not have resulted in as strong of results due to the differences in intervention delivery relative to Bryant and colleague's original study (2011). First, this study was completed 3 days per week for 45 minutes per session across 22 weeks. Six of the units from ENI1 were completed. In the Bryant et al. (2011) study where significant gains were seen for participants, sessions occurred 4 days per week for 25 minutes per session across 19 weeks and all 10 units were completed. In the Bryant et al. (2008) study, 64 sessions occurred for 15 minutes per session across 18 weeks. During the Bryant et al. (2008) study, they did not support ENI1 and hypothesized that the intervention was not delivered frequently enough. Therefore, it is important to consider intervention intensity in decision-making regarding interventions. While ENI1 did not occur as often in this study compared to the Bryant et al. (2011) study, the length of the interventions was very similar with this study being 3 weeks longer.

While the intensity of the intervention describes how frequent the intervention occurs, the length of intervention describes how many weeks an intervention is put into place. Another reason the study did not see as strong results may be due to the longer length of the intervention. Smaller effects have been found with early numeracy interventions that last longer than 8 weeks (Nelson and McMaster, 2019). During this study, the intervention was implemented for 20 weeks, which could explain the smaller effect sizes that were found. Additionally, this could potentially explain why participants started scoring in acceptable ranges on the Pre-Tests of the Unit Checks during later Units. Early numeracy interventions may show larger effects if intervention sessions are more frequent over a shorter intervention length.

Measurement of Outcomes. Another difference with the current study is that in the single case design, curriculum-based measurements were used as dependent variables to assess intervention outcomes. Former studies have used the Texas Early Mathematics Inventories – Progress Monitoring measure, the Texas Early Mathematics Inventories – Outcomes measure, and the Stanford Achievement Test – Tenth Edition (Bryant et al., 2011). One consideration for the current study is when certain early numeracy skills were introduced in ENI1 and the dependent variables used to assess outcomes. For example, lessons related to Place Value began during week four of the intervention, which was Unit 2, Lesson 4 (Appendix C). When the skill was introduced in the intervention, four out of five participants performed better on subsequent assessments of Place Value. This may be why an effect was less noticeable in the first AB phase for Place Value. The concept of decomposing was first introduced during Unit 1 Lesson 9 with numbers within 10. Although the dependent measures used in this study assessed important early numeracy skills, they targeted specific discrete skills. Whereas, the outcome measures for previous ENI1 studies assessed more global early numeracy skills.

A related issue for the evaluation of intervention outcomes with Decomposing is also evident in the data. Similar to Place Value, Decomposing was addressed in the intervention until the middle of Intervention Phase 1. Also, ENI1 worked with decomposing numbers within 10. Decomposing tested the skills with numbers as large as 20. When the skills were introduced, increases in performance were not seen in the measure of Decomposing. This may have been due to higher numbers being used in the measure versus what was taught in the intervention.

A last consideration for the dependent variables used in this study and the ability to evaluate intervention outcomes with them has to do with their sensitivity to growth when used for frequent progress monitoring. For all participants, PV and DC had relatively flatter trends when compared to Number Identification. Average weekly growth from fall to spring for Place Value is 0.08 points per week and Decomposing is 0.15 points per week (FastBridge Learning, LLC 2020). Therefore, it would be expected that a student would take 13 weeks to get one more point in Place Value and 6 weeks to get one more point in Decomposing. This might explain why the growth may not have been apparently seen in the data.

Academic Engagement. While formal data on academic engagement was not collected, informal observations during assessment suggested that students who showed better academic engagement during the intervention showed stronger outcomes. Two participants, Olive and Polly, had significant behavioral challenges throughout the intervention. Behaviors of concern include putting their head down, crying, sitting at another table, pushing away the work, and refusal to complete the work. Conclusions about why these behaviors occurred cannot be drawn from the results. These two participants did not make great growth during the intervention. Participant characteristics can play a role in facilitating or discouraging intervention integrity. When challenging behavior occurs, it may make it more difficult for educators to implement the intervention for those students and may hinder the engagement of those students in the intervention (Roach, Lawton, & Elliott, 2014). Potentially this intervention may be too traditional to use in a Montessori setting (i.e., more sitting and working rather than the open work environment).

Limitations

As discussed, there were limitations to the current study. First, the decision-making about phases of the study were a limitation, including when to move from baseline to intervention, when phases were conducted, and the replication phase of the study. A second area of limitations relates to features of the dependent variables selected for the study. A final limitation is the behavioral challenges that occurred with two of the participants and the degree to which participant behaviors negatively impacted intervention integrity.

Experimental Design. This study did not use visual analysis of the data to decide when to move from a stable baseline to intervention. Further, visual analysis was not used to determine when to introduce the withdrawal phase. Steady trends in baseline data allow researchers to be able to predict future behavior if nothing were to change (Riley-Tillman & Burns, 2009). Four out of five participants showed variable data in at least one measure during the baseline phases; therefore, interpretations of causality related to the introduction of the intervention are more challenging.

Rather than relying on trends in the data to make decisions about phase changes, the baseline and withdrawal phases were matched to natural breaks in the school's calendar. This introduced some additional confounds. The study did not begin until mid-fall, therefore, the Intervention 1 phase was much shorter than the Intervention 2 phase where growth was seen. It may also be possible that early numeracy concepts are skills that need consistent practice to build up over time. Therefore, it may be possible that there was not enough time and data in the Intervention 1 phase for a pattern to emerge in the data or for a trend to occur. This is particularly evident for Lola, Nova, and Olive's data.

An ABAB design is meant to control for outside influences on the outcome being evaluated, because there is an assumption that the only thing that changes during the A phases is the presence of the intervention or not (Riley-Tillman and Burns, 2009). This is accounted for in the first AB phases, but then the replication of the effect is confounded by the fact that part of the second A phase was the holiday break. Therefore, classroom instruction was also removed during this phase of the intervention.

Lastly, during the withdrawal phase, of an ABAB design, the data should return to similar scores of the first baseline phases (Riley-Tillman and Burns, 2009). This would allow assumptions to be made that the intervention was the impacting factor to the wanted outcomes. However, with academic interventions, like ENI1, during the withdrawal phase the data did not return to baseline. This is presumed because when an academic skill is taught to the participants, they learned how to complete that skill consistently, rather than losing the academic skill. That is seen in the participants data in this study. Many participants slightly decreased during the withdrawal phase but did not fall completely back to the first baseline phase like is typically expected with an ABAB design.

Dependent Variables. The dependent variables chosen for this study may not have been sensitive enough to have shown growth during the short period of the phases, specifically the first intervention phase. Typical rates of improvement are smaller for the dependent variables chosen (Place Value is 0.08 points per week from fall to spring; and Decomposing is 0.46 points per week from fall to spring). Growth that may have occurred may not have been captured in the data. This could explain why participants' Place Value and Decomposing scores appeared flatter especially for Place Value where growth was seen when looking at the mean scores. If the first intervention phase was longer, more growth may have potentially been seen, similar to the second intervention phase. Future research should explore the use of other early numeracy measures that are more sensitive to smaller changes in skills.

Behavioral Challenges. Finally, another variable to consider was participants' academic engagement during intervention, specifically Olive and Polly. Olive and Polly did not show great growth throughout the intervention. This may have been impacted by their lack of academic engagement especially when behavioral needs including putting their head down, sitting at other tables, and crying were occurring. When these challenges occurred, their academic engagement was low. When behavior challenges did not occur, their academic engagement was observed to be higher and they participated throughout the intervention session. Thus, when behavior was a concern, they may not have received the full instruction for that session, which could have impacted their performance overall. These behavioral challenges make it difficult to draw conclusions around ENI1 and the participants' performance who had behavioral challenges.

Implications for Practice

The current study has findings that are important for educators who work on early numeracy concepts with students in nontraditional school settings. The ENI1 intervention shows promise when used with students in a Montessori setting. The classroom teacher and participants found this intervention to be acceptable. The approach of the intervention, which is more direct and explicit, is different from the Montessori philosophy. Therefore, teacher acceptability of ENI1 is important as the approach of instruction is different but is still found to be acceptable to use within a Montessori setting.

Promoting academic engagement is important to consider to make sure participants were engaged. This study showed that the intervention was delivered with fidelity, but participant behavior may suggest that a couple participants did not participate with fidelity. Since two participants struggled more to participate, it may be important to consider how to teach and reinforce desired behaviors of participation to make sure participants who are used to a more exploratory learning environment are supported as they participate in a more structured intervention. Further, the intervention was easily implemented by individuals that were not the original researchers or classroom teachers. This shows the transportability of the ENI1 intervention and it shows that interventionists do not need to have high levels of training.

This study also highlights the importance of selecting progress monitoring tools that are well matched to skills targeted within an intervention and are sensitive to smaller changes in skills. Shapiro (2011) identifies general outcome measures has a standardized assessment tool across curricular objectives and is designed to be sensitive enough to instruction gains across time. A subskill mastery measure assesses specific skills being taught. When measuring progress toward a long-term goal, general outcome measures are more likely to show progress than the use of subskill mastery measures (Shapiro, 2011).

The match of the progress monitoring assessments and the intervention targets needs to be considered. For example, Number Identification seems to be sensitive to many forms of math practice, because all math practice requires repeated practice in reading numbers to solve other types of problems. Another impact of using measures with limited sensitivity would be lower effect sizes. If scores do not change, there is less change that the effect size will look good.

Last, intensity of intervention as related to frequency per week is important to consider when evidence-based interventions are adopted. During the Bryant et al. (2008) study, the researchers discussed that the lack of progress may have occurred because of a lack of frequency of intervention per week. Therefore, during the Bryant et al. (2011) study, the researchers increased the frequency of the intervention during the week. The researchers found gains in progress during that study. Therefore, more frequent practice of the skills taught in the intervention throughout the week may lead to an increase in progress of those skills.

Future Research

Future single case design research should look at early numeracy measures that are more sensitive to changes in skills. A general outcome measure for early numeracy may show more increments of growth for the participants during a shorter intervention period. Second, continued research on ENI1 and the optimal frequency for weekly intervention sessions is suggested. Third, future research should consider the importance of planning for and collecting data on participant engagement as an indicator of intervention integrity. Also, research should examine whether ENI1 is appropriate for students with behavior challenges. Formally assessing academic engagement may help determine the effects for students with behavioral challenges. Considering how to teach and reinforce desired behaviors of participation to make sure students who are used to more exploratory learning environments are supported in the intervention will be important. Lastly, examining the effects of ENI1 with nontraditional students will determine whether transportability can occur in a nontraditional educational setting. Overall, results show promise, but more research is needed.

References

Baglici, S. P., Codding, R., & Tryon, G. (2010). Extending the research on the tests of early numeracy: Longitudinal analyses over two school years. *Assessment for Effective Intervention, 35*(2), 89-102. doi:http://dx.doi.org/10.1177/1534508409346053

Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N., & Chavez, M. M. (2008). Mathematics intervention for first- and second- grade students with mathematics difficulties: The effects of tier 2 intervention delivered as booster lessons. *Remedial & Special Education, 29*(2), 20-32.

Bryant, D. P., Bryant, B.R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematical difficulties. *Exceptional Children, 78*(1), 7-23.

Christ, T. J., and Colleagues. (2017). earlyMath time-limited measures observing & rating administrator accuracy. FastBridge Learning, NCS Pearson, Inc.

Codding, R. S., Chan-Iannetta, L., George, S., Ferreira, K., & Volpe, R. (2011). Early number skills: Examining the effects of class-wide interventions on kindergarten performance. *School Psychology Quarterly, 26*(1), 85-96. doi:http://dx.doi.org/10.1037/a0022661

Common Core State Standards Initiative Preparing America's Students for College & Career. (2019). Mathematics Standards. Retrieved from http://www.corestandards.org/Math/.

Conoyer, S. J., Foegen, A., & Lembke, E. S. (2016). Early numeracy indicators. *Remedial & Special Education, 37*(3), 159-171. doi:10.1177/07419325619758

Dyson, N. I., Jordan, N. C., & Glutting, J. (2011). A number sense intervention for low-income kindergarteners at risk for mathematic difficulties. *Journal of Learning Disabilities, 46*(2), 166-181. doi: 10.1177/0022219411410233

FastBridge Learning, LLC (2020). Retrieved from: http://www.fastbridge.org/

Fuchs, L. S., Fuchs, D., & Karns, K. (2001). Enhancing kindergarteners' mathematical development: Effects of peer-assisted learning strategies. *The Elementary School Journal, 101,* 495–585.

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice (Wiley-Blackwell), 22*(1), 36-46. doi:10.1111/j.1540-5826.2007.00229x

Kiss, A. J., Nelson, G., & Christ, T. J. (2019) Predicting third-grade mathematics achievement: A longitudinal investigation of the role of early numeracy skills. *Learning Disability Quarterly, 42*(3), 161-174.

Lembke, E., & Foegen, A. (2009). Identifying early numeracy indicators for kindergarten and first-grade students. *Learning Disabilities Research & Practice, 24*(1), 12-20.

Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41*(5), 451-459.

Mononen, R., Aunio, P., Koponen, T., & Aro, M (2014). A review of early numeracy interventions for children at risk in mathematics. *International Journal of Early Childhood Special Education, 6*(1), 25-54.

National Assessment of Educational Progress (NAEP). (2015). The nation's report card. https://www.nationsreportcard.gov/reading_math_2015/#?grade=4

National Center on Intensive Intervention at American Institutes for Research (2019). Principles for designing intervention in mathematics. U.S. Department of Education. Washington DC.

NCS Pearson, Inc. (2020). *aimsweb Technical Manual.* Bloomington, MN: NCS Pearson, Inc.

Nelson, G., & McMaster K. L. (2019). The effects of early numeracy interventions for students in preschool and early elementary: A meta-analysis. *Journal of Educational Psychology, 111*(6), 1001-1022. doi: 10.1037/edu0000334

Riley-Tillman, T. C., & Burns, M. K. (2009). *Evaluating educational interventions: Single-case design for measuring response to intervention*. The Guilford Press.

Roach, A. T., Lawton, K., & Elliott, S. N. (2014). Best practices in facilitating and evaluating the integrity of school-based interventions. In P. L. Harrison & A. Thomas (Eds.), Best practices in school psychology: Data-based and collaborative decision making. Bethesda, MD: National Association of School Psychologists.

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52*(9), 1190-1197.

Shanley, L., Clarke, B., Doabler, C. T., Kurtz-Nelson, E., and Fien, H. (2017). Early number skills gains and mathematics achievement: Intervening to establish successful early mathematics trajectories. *The Journal of Special Education, 51*(3), 177-188. doi: 10.1177/0022466917720455

Shapiro, E. S. (2011). *Academic skills problems fourth edition: Direct assessment and intervention.* The Guilford Press.

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnston IV Tests of Achievement.* Rolling Meadows, IL: Riverside.

Toll, S. M., & Van Luit, J. H. (2012). Early numeracy intervention for low-performing kindergartners. *Journal of Early Intervention, 34*(4), 243-264. doi:10.1177/1053815113477205

Theodore J. Christ and Colleagues (2018). Formative Assessment System for Teachers™ Technical Manual, Minneapolis, MN: Author and FastBridge Learning.

Van Luit, J. E. H. & Van de Rijt, B. A. M. (2009). Early Numeracy Test – Revised. Doetinchem, Netherlands. Graviant.

What Works Clearinghouse. (2009). *Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools.* Retrieved from https://ies.ed.gov/ncee/wwc/practiceguide/2.

What Works Clearinghouse. (2013). *Teaching Math to Young Children.* Retrieved from https://ies.ed.gov/ncee/wwc/PracticeGuide/18.

Appendix A

Teacher Social Acceptability Survey

Hello, my name is Melissa Trites. The purpose of this survey is to get your perceptions of how acceptable the goals, procedures, and outcomes were for the first grade Early Numeracy Intervention Level 1 program.

Thank you for allowing your student(s) to be a part of the UWEC Human Development Center Academic Intervention Clinic math program. The Early Numeracy Intervention Level 1 (ENI1) was delivered three times per week for about 40 minutes each day. ENI1 targeted number sequencing, number knowledge and relationships (i.e., comparing numbers and quantity of numbers and ordering numbers), and place value. The ENI1 program also included activities that targeted a beginning understanding of addition and subtraction. Each day, the intervention included modeling from the lead interventionist, guided practice with feedback from the interventionist, and independent practice.

Social Acceptability for Teacher

1. This is an acceptable intervention for students at a Montessori elementary school.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

2. Most Montessori teachers would find this intervention appropriate.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

3. I would suggest the use of this intervention to other Montessori teachers.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

4. The intervention is appropriate to meet a Montessori school's needs and mission.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

5. Most Montessori teachers would find this intervention suitable for the described purposes and mission.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

6. This intervention would not result in negative side-effects for the students.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

7. This intervention is consistent with those I have used or are offered in this school setting.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

8. The intervention is a fair way to fulfill the intervention purposes.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

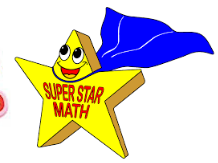
9. This intervention is reasonable to meet the stated purposes.

   Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

10. Overall, this intervention would be beneficial for a variety of Montessori elementary school students.

    Strongly disagree   Disagree   Neutral   Agree   Strongly Agree

Participant Social Acceptability Survey

What did you think?

	1 = not true at all	**2 =** not true	**3 =** true	**4 =** very true
1. The math practice I did this year was good for me.	1	2	3	4
2. I liked the way I practiced math this year.	1	2	3	4
3. I think my math skills improved this year.	1	2	3	4
4. I think other kids would think this math program is helpful.	1	2	3	4
5. I would want to work on math with university students again.	1	2	3	4
6. This program helped me understand math better.	1	2	3	4
7. Overall, I think the math work I did here was helpful.	1	2	3	4

What else would you like to say about our math program?

____

____

Appendix B

Procedural Integrity for an Intervention Lesson

Teacher:____________________________School:_____________________

Reliability completed by:_______________Percentages:_________________

✓ Behavior Demonstrated
(Blank) Behavior Not Demonstrated
NA Not Applicable

Lesson
___ **1.** Teacher reviews previous knowledge verbatim.
___ **2.** Teacher writes before, after, next to on their wipe board.
___ **3.** Teacher gives each student a number path.
___ **4.** Teacher explains vocabulary: “Before”.
___ **5.** Teacher writes 3 with a blank before it.
___ **6.** Teacher has students find what comes before 3 on the number path.
___ **7.** Teacher explains vocabulary: “after”.
___ **8.** Teacher writes 6 with a blank after it.
___ **9.** Teacher has students find what comes after 6 on the number path.
___ **10.** Teacher explains vocabulary: “next to”.
___ **11.** Teacher writes an 8 with a blank before and after it.
___ **12.** Teacher has students find what numbers are next to 8 on the number path.
___ **13.** Teacher writes a 12 with a blank after it.
___ **14.** Teacher asks the students what number it is and what number is missing.
___ **15.** Teacher asks what comes after, before, and what numbers are next to 12.
Guided Practice
___ **16.** Teacher has student select number cards.
___ **17.** Teacher has students circle the numbers that are before, after, or next to the numbers chosen.
___ **18.** Teacher circles number 16 on their number path and writes the number 14 with a blank after on the wipe board.
___ **19.** Teacher has students look at both the number path and wipe board.
___ **20.** Teacher asks if 16 is the number after 14 and why.
___ **21.** Teacher gives the guided practice sheet to students.
___ **22.** Teacher prompts and supports students during guided practice sheet.
Independent Practice

___ **23.** Teacher has students complete as many problems as they can in one minute.
___ **24.** Teacher goes over the answers and has students correct their work with a different color on their own sheets.

Appendix C

Progression of Skills Targeted in Units and Lessons

Unit/Lesson	Skills Targeted
Unit 1	
Lesson 1	Finding neighbor numbers
Lesson 2	Putting numbers in groups
Lesson 3	Learning parts and wholes of numbers, idea of adding and subtracting
Lesson 4	Missing number in sequences of 3, using numbers up to 40
Lesson 5	Putting numbers in groups and counting groups
Lesson 6	Breaking numbers apart, i.e., 4 = 2 and 2, (1-4), writing as equations
Lesson 7	Ordering numbers least to greatest up to 12
Lesson 8	Putting numbers into groups of 10 and counting them
Lesson 9	Decomposing numbers 5-7 into parts
Lesson 10	Identifying missing numbers in 3 number sequences, using numbers up to 55
Lesson 11	Identifying groups of 10 and how many in all by counting by 10s
Lesson 12	Decomposing numbers 8-9 into parts
Unit 2	
Lesson 1	Counting two groups and deciding if the groups are equal
Lesson 2	Finding neighbors of numbers, i.e., what number comes before and after 36
Lesson 3	Counting, adding, deciding if the equation equals a number, and making equations.
Lesson 4	Counting groups of 10 and some leftover ones, i.e., counting rods and ones
Lesson 5	Finding neighbors of numbers, before and after numbers, using numbers up to 50
Lesson 6	Counting groups of 10 and some leftover ones, i.e., counting rods and ones
Lesson 7	Learning the first 2 steps of the Math Scene Investigator (MSI) strategy for word problems, step 1 is to inspect and find clues, step 2 is to plan and solve
Lesson 8	Learning partner equations, i.e., 1+3=2+2
Lesson 9	Skip counting by 10s
Lesson 10	Learning partner equations, i.e., 1+4=3+2
Lesson 11	Comparing numbers and deciding which number is less than the other
Lesson 12	Identifying missing numbers in a three-number sequence, using numbers up to 40

Lesson 13	Ordering number from least to greatest, using numbers up to 40
Lesson 14	Reviewing the steps of the MSI strategy and then learning the last step, the last step is retrace
Unit 3	
Lesson 1	Adding and subtracting 1 and 0
Lesson 2	Identifying numbers that are before, after, or next to a number using a hundred chart
Lesson 3	Identifying the symbol and then adding or subtracting 0 or 1
Lesson 4	Counting rods and cubes to make numbers
Lesson 5	Identifying a missing number in a 3-number sequence by counting up or counting down
Lesson 6	Counting rods and cubes to make numbers
Lesson 7	Solving addition and subtraction word problems using the MSI strategy
Lesson 8	Using addition and subtraction to solve neighbor minus neighbor (i.e., 9-8) problems and number minus number (i.e., 9-9) problems
Lesson 9	Skip counting by 5s
Lesson 10	Using addition and subtraction to solve neighbor minus neighbor problems and number minus number problems
Lesson 11	Identifying what number is greater between two numbers
Lesson 12	Identifying a missing number in a 3-number sequence by counting up or down, using numbers up to 50
Lesson 13	Reading a bar graph and writing the numbers from the graph, comparing two numbers to identify which is less
Lesson 14	Solving addition and subtraction word problems using the MSI strategy
Unit 4	
Lesson 1	Adding and subtracting 1, 2, or 3 to or from a number
Lesson 2	Identifying number before, after, or next to another number on a hundred chart, then writing those numbers
Lesson 3	Adding or subtracting 1, 2, or 3 to or from a number
Lesson 4	Counting rods and cubes to make a number
Lesson 5	Identifying a missing number in a 3-number sequence by counting up or down, crossing decade numbers (i.e., 10, 20)
Lesson 6	Adding and subtracting groups of 10 using a hundred chart
Lesson 7	Solving addition and subtraction word problems using the MSI strategy
Lesson 8	Adding and subtracting 1, 2, or 3 to or from a number
Lesson 9	Skip counting by 5s from 0 to 100
Lesson 10	Adding and subtracting 1, 2, or 3 to or from a number using a number path
Lesson 11	Writing 1- and 2-digit numbers and then identifying the number that is lesser of 2 numbers

Lesson 12	Identifying a missing number in a 3-number sequence by counting up or down, using numbers up to 55
Lesson 13	Reading bar graphs, using the numbers to order from least to greatest
Lesson 14	Solving addition and subtraction word problems using the MSI strategy
Unit 5	
Lesson 1	Adding and subtracting 3 from numbers using a part-part-whole mat
Lesson 2	Identifying missing numbers on a hundred chart
Lesson 3	Adding and subtracting 1, 2, or 3 from a number
Lesson 4	Counting rods and cubes to make numbers, making numbers up to 60
Lesson 5	Identifying a missing number in a 3-number sequence by counting up or down, crossing decade numbers, using numbers up to 50
Lesson 6	Adding and subtracting groups of 10 from a number using a hundred chart
Lesson 7	Solving addition and subtraction word problems using the MSI strategy
Lesson 8	Adding and subtracting 1, 2, or 3 from a number and filling out the fact families (i.e., 5+4=9, 4+5=9, 9-5=4, 9-4=5)
Lesson 9	Skip counting by 2s from 0-100
Lesson 10	Adding and subtracting 1, 2, or 3 from a number and filling out the fact families
Lesson 11	Identifying the number that is greater of two numbers
Lesson 12	Identifying a missing number in a 3-number sequence by counting up or down, using numbers up to 55
Lesson 13	Ordering numbers from greatest to least, using numbers up to 50
Lesson 14	Solving addition and subtraction word problems using the MSI strategy
Unit 6	
Lesson 1	Identifying double facts, solving addition double facts, writing the related facts (i.e., 7+7=14, 14-7=7)
Lesson 2	Identifying number before, after, and next to another number on a number path
Lesson 3	Identifying double facts, solving addition double facts, writing the related fact
Lesson 4	Counting rods and cubes to make numbers, using numbers up to 90
Lesson 5	Identifying a missing number in a 3-number sequence by counting up or down, crossing decade numbers, using numbers up to 95
Lesson 6	Adding and subtracting groups of 10 from a number using a hundred chart
Lesson 7	Solving addition and subtraction word problems using the MSI strategy
Lesson 8	Identifying double facts, solving addition doubles facts by counting by 2s
Lesson 9	Skip counting by 2s from 2 to 100

Lesson 10	Solving addition double facts by using the count by 2s strategy and a number path
Lesson 11	Identifying the lesser number of two numbers
Lesson 12	Identifying a missing number in a 3-number sequence by counting up or down, using numbers up to 100
Lesson 13	Counting rods and cubes to make numbers, using numbers up to 50
Lesson 14	Solving addition and subtraction word problems using the MSI strategy