Tier 3 Intensified Intervention for Second Grade Students with Severe Mathematics Difficulties

Diane Pedrotty Bryant\textsuperscript{1,2,*}, Brian R. Bryant\textsuperscript{1,2}, Danielle A. Sorelle-Miner\textsuperscript{1}, Terry S. Falcomata\textsuperscript{2}, Maryam Nozari\textsuperscript{1,2}

Authors’ affiliations:
\textsuperscript{1} The Meadows Center for Preventing Educational Risk
\textsuperscript{2} Department of Special Education, The University of Texas at Austin

* Corresponding author: E-mail: dpbryant@austin.utexas.edu; Phone: 512-784-7346

Authors’ e-mail addresses:
Brian R. Bryant: brianbryant@aol.com
Danielle A. Sorelle-Miner: dsorelle@austin.utexas.edu
Terry S. Falcomata: falcomata@austin.utexas.edu
Maryam Nozari: Maryam.Nozari@utexas.edu

Abstract

A multiple baseline design was employed to examine the effects of an intensive mathematics intervention, which focused on early numeracy concepts and skills. Thirty-three second grade students participated in the study. Students attended five different schools in one school district and received the intervention in a total of 12 groups from mathematics interventionists. The intervention occurred for about 30 min per session, five days a week for eight weeks with a maintenance phase two weeks later and generalization testing four weeks later. The intervention consisted of explicit instruction, strategies, and mathematics practices. Visual analyses of the data showed improvement for the majority of the groups. Effect size calculations showed no evidence of overlapping data between baseline and intervention for nine out of twelve groups. Maintenance data revealed positive results and five students had a posttest score at or above the 25\textsuperscript{th} percentile on the generalization measure. We discuss implications for practice with limitations and future research.

Keywords: early numeracy, intensive mathematics intervention, number, operation, algebraic thinking, severe mathematics difficulties, tertiary intervention, Tier 3

A multi-tiered framework has been employed widely in Response to Intervention (RtI) models since 2000 and is a vital component of the Multi-tier Systems of Support (MTSS) framework, as well. In recent years, there has been an increase in the number of studies focused on investigating the effects of interventions for improving the mathematics performance of elementary level students with mathematics difficulties (e.g., Clarke et al., 2016; Dyson, Jordan, & Glutting, 2013; Fuchs, Geary, Fuchs,
Compton, & Hamlett, 2014). Even so, researchers have demonstrated that about 5% to 8% of the school-aged population experiences persistently low mathematics achievement and thus requires more intensive intervention (Berch & Mazzocco, 2007; Geary, 2011; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Murphy, Mazzocco, Hanich & Early, 2007), with many of these struggling students being identified as having a mathematics learning disabilities (MLD) (Murphy et al., 2007). Moreover, findings from the National Assessment of Education Progress (NAEP), 2015 showed that students with mathematics difficulties and disabilities demonstrated the greatest lags in performance. Although these findings are not particularly surprising, regarding students with all disabilities, LD has long been the largest category of disabilities, so it makes sense to infer from these data that the majority of the students with disabilities are students with LD. For example, fourth-grade findings for students who are chronically low achievers in mathematics showed scores below the Basic Level of performance on the mathematics assessment; 46% of students with disabilities scored below basic compared to 15% of students without disabilities. Additionally, low-income students are disadvantaged (28% falling below the Basic Level) compared to students from higher economic status (8%), which is a pattern that persists throughout the school years (Murnane, & Duncan, 2011; Reardon, 2011). Together, these national assessment findings are concerning and indicate that it is imperative to understand the mathematics difficulties students exhibit for intensive interventions to be developed and tested for teaching the mathematics curriculum.

**Students with Severe Mathematics Difficulties and Mathematics Curriculum**

Regarding mathematics difficulties, we defined severe difficulties as students with mathematics achievement scores that fell at or below the 10th percentile (Mazzocco Feigenson, Halberda, 2011) as measured by the school district’s universal screener and a proximal measure (Bryant et al., 2008). This group of elementary level students with identified severe mathematics difficulties typically manifest problems with the concepts and skills of early numeracy (e.g., counting [e.g., 10, 11, 12, 13, 14], comparing the magnitude of numbers [46 > 39], identifying the missing number in a sequence of numbers on a number line [e.g., ___, 15, 16 = 14], 15, 16]), the base 10 numeration system (e.g., understanding the place and value of digits in numbers), and arithmetic (e.g., composing [e.g., combining two addends, 4 + 3 = 7], decomposing [6 + 7 = 6 + 6 + 1 = 13], retrieving solutions automatically) number combinations or basic facts (Jordan, Glutting, Dyson, Hassinger, & Irwin, 2012; Geary et al., 2009).

Butterworth and Reigosa (2007) and Geary (2011) theorized that developmental delays in number systems (e.g., understanding numerical magnitude) contribute to difficulties with comprehension of the base-10 system and automatic fact retrieval. For example, students in second grade should be able to group base-10 models into 10s (e.g., 10 ones = 1 ten) and 100s (e.g., 10 tens = 100) and ungroup a bundle of 10 into 10 ones and a bundle of 100 into 10 tens as a prerequisite skill for computing problems with regrouping such as 42 + 89 = 131 and 23 = 102 – 79 (Fuson, 2012; Van de Walle, Karp, & Bay-Williams, 2016). Yet, for many students with mathematics difficulties, the idea that the place in which digits occur in a numeral represents a power of 10 value, for instance, can be abstract and elusive. Automatic fact retrieval is another challenge due to inadequate and delayed understanding of number and numeration which in turn affects students’ abilities to successfully use counting strategies to
solve addition problems (e.g., counting-on-from-first strategy, $2 + 5 = ?$, “2 [pause], 3, 4, 5, 6, 7” – the solution is 7; counting-on-from-larger strategy or the min strategy, $2 + 5 = ?$, “5 [pause], 6, 7 - the solution is 7) and counting strategies to solve subtraction problems (e.g., counting up from given strategy, $5 + ? = 7$, “5 [pause], 6, 7- the solution is 2 (Carpenter & Moser, 1984; Geary, 2004, 2011). Although students with severe mathematics difficulties are capable of recalling number combinations, they recall fewer facts and make more errors compared to their typically achieving peers, which in turn effect calculation solutions for two- + two-digit or three- + three-digit problems (Andersson, 2010; Geary, Hoard, Nugent & Bailey, 2012). Difficulties with fact retrieval have been theorized to stem from working memory problems associated with an inability to inhibit unrelated information when students are trying to recall facts (Geary, 2011). Unfortunately, the long-term ramification is difficulties with fact retrieval, which can affect students’ performance when solving whole number computations and word problems (L. Fuchs et al., 2005).

Improving mathematics outcomes is critical for those students who manifest severe mathematics difficulties. A rigorous mathematics curriculum is necessary to increase the mathematics performance of all students to be competitive in the global workforce and succeed in postsecondary education (National Mathematics Advisory Panel, ([NMAP], 2008).

The Common Core State Standards for Mathematics ([CCSSM]; NGACBP, CCSSO, 2010) can influence the development of a rigorous mathematics curriculum for school districts. Although not a curriculum per se, the standards were developed to provide more coherence and focus for mathematics curricula, which school districts and states have in place, on critical mathematical ideas. The CCSSM were based on the notion that mathematics is hierarchical with specific learning trajectories of concepts and skills for mathematics domains (e.g., number and operations in base ten, operations and algebraic thinking). Learning trajectories are developmental progressions that focus on specific “big ideas” (e.g., number and operation – counting, comparing and ordering numbers, equivalence) related to each domain (Clements & Sarama, 2014). Students who fail to learn early critical mathematics concepts and skills associated with the developmental progressions in the primary grades are likely to continue to struggle as they progress through later grades and are required to master increasingly complex and demanding mathematics standards and curriculum (Bryant et al., 2008; Geary, 2004, 2011; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Morgan, Farkas, & Wu, 2009). Thus, it is imperative for young children with severe mathematics difficulties to receive intensive interventions to improve their mathematics performance.

**Intensifying Interventions for Students with Severe Mathematics Difficulties**

Considerable work has been done demonstrating the effectiveness of supplemental mathematics interventions (i.e., Response to Intervention: Tier 2) with elementary level students with mathematics difficulties (B. Bryant et al., 2014; Clarke et al., 2014; Dyson, Jordan, Beliaikoff, & Hassinger-Das, 2015; Gersten et al., 2015; Strand Cary, Clarke, Doabler, Smolkowski, Fien, & Baker, 2017; Powell et al., 2015). For example, Gersten et al. (2015) conducted a replication study of Fuchs et al. with 994 at-risk students randomly assigned to treatment or control conditions. Students in the intervention condition received instruction for about 30 hours of small-group work and classroom instruction; whereas, students in the control condition received typical instruction or “business as usual”
with teacher assistance. Results showed that students in the intervention condition showed significantly superior performance on a distal measure of mathematics achievement. Although these findings are encouraging, there remains a group of students who repeatedly demonstrate inadequate response to empirically validated secondary (Tier 2) interventions. Thus, intensified interventions are warranted, which build upon the strengths of interventions that prove important for struggling, Tier 2 students (Bryant et al., 2011; Dyson et al., 2013; L. Fuchs et al., 2008; Fuchs, Fuchs, & Malone, 2018). Conceivably, Tier 3 or tertiary intervention is necessary for high-risk or persistently low performing students with severe mathematics difficulties who require more intensive interventions, which may or may not include special education services depending on how individual states view Tier 3 services (Ludlow, 2014).

Typically, in a multi-tiered RTI approach, students receive Tier 3 intensified intervention after they have demonstrated inadequate response to one or two iterations of Tier 2, supplemental instruction; however, this is not always the case. According to Ervin (2015),

there should be a mechanism through which students who are experiencing very severe or significant academic, behavioral, or social-emotional problems can be triaged directly into Tier 3 to receive necessary intensive and individualized intervention supports. For some students, the second option is necessary to provide needed supports in a timely fashion rather than delaying access to these supports by making students wait to go through Tier 1 and Tier 2 intervention services (http://www.rtinetwork.org/essential/tier3/consideringtier3).

For purposes of this study, participants were identified through universal screening procedures (see Participants section) as having severe mathematics difficulties; thus, they were eligible for Tier 3 intensive intervention. Given that mathematics difficulties can be resistant to change in later grades due in part to the hierarchical learning progression of the mathematics curriculum, early intensive intervention is warranted (Doabler et al., 2015; Geary et al., 2011).

The Tier 3 intensified intervention employed in this study was grounded in instructional design that focused on a conceptual framework for primary-aged students with mathematics difficulties based on explicit, systematic instruction of an early numeracy curriculum combined with cognitive strategies (Swanson, Hoskyn, & Lee, 1999). Skills and concepts students should know at the end of second grade, include having a strong foundation in early numeracy and numeration (e.g., relationship between quantity and unit, quantity discrimination, equality, counting, place value), understanding the meaning and properties of addition (i.e., commutative, associative), recognizing the inverse relationship of addition with subtraction, applying efficient counting strategies, being able to retrieve basic facts, and understanding place value, (Caldwell, Karp, & Bay-Williams, 2011; Dougherty, Flores, Louis, & Sophian, 2010). In short, students should have a solid foundation in “number sense,” which is defined as “moving from the initial development of basic counting techniques to more sophisticated understandings of the size of numbers, number relationships, patterns, operations, and place value” (National Council of Teachers of Mathematics [NCTM], 2000, p.79).

To intensify instruction, we borrowed from Vaughn, Wanzek, Murray, and Roberts’ (2012) four considerations for intensifying instruction. First, research findings for “in-
tegrating strategies that support cognitive processing through academic instruction may accelerate academic progress” (p. 7). Cognitive processing difficulties associated with semantic memory (theoretical framework includes a portion of long-term memory, which is responsible for storing information such as facts, meanings, concepts and knowledge, McLeod, 2010) and working memory (theoretical framework of structures and processes used for the temporary storage and manipulation of information including executive functioning, McLeod, 2012) have been linked to problems with, for example, basic facts in terms of accurately committing to and retrieving them from long-term memory (Geary, 2012; Swanson & Jerman, 2006). However, rather than training these cognitive processes separate from academic instruction, the focus should be on teaching “memory-enhancing techniques” (p. 15) such as the use of mnemonics, verbal rehearsal, graphic organizers, and cognitive strategies (Vaughn et al., 2012). In the current study, intensified instruction was conducted, in part, by having the mathematics interventionists “think-aloud” as they solved problems and providing specific praise, such as, “You used the Make Ten Plus More Strategy to solve 9 + 4 = .” to help students remember the content. The interventionists also focused their students’ attention on self-regulating (e.g., self-talking through the steps for solving a word problem) their use of strategies for solving various problems.

Second, the mathematics interventionist delivers intensified instruction that is explicit and systematic. In a classic meta-analysis of intervention treatment outcomes, Swanson et al. (1999) found explicit, systematic instructional delivery combined with cognitive strategy training as having high effect sizes on the performance of students with learning disabilities (LD). Coyne, Kame’enui, and Carnine (2011) identified specific evidence-based principles of explicit instruction including focusing on important content, making learning strategies transparent to the learner (e.g., for the basic facts), scaffolding instruction, checking for understanding, and providing review opportunities for designing interventions. In this study, the mathematics interventionist intensifies explicit, systematic instruction by providing more modeling, increasing practice opportunities, providing more scaffolded instruction, and using multiple representations (e.g., number lines, ten frames) to visualize the mathematics. Vaughn et al. (2012) also identified systematic instruction as a necessary component for intensified instruction. Mathematics skills were broken down into smaller skills and carefully sequenced during systematic instruction; therefore, controlling for task difficulty and reducing “cognitive load,” as a means for optimizing cognitive capacity for learning (Sweller, 2005).

Third, the amount of instructional time should be examined to determine whether it is sufficient for educators to provide students with multiple opportunities to practice with corrective feedback. In the current study, the interventionist delivered instruction five days per week for the duration of the study, which is an increase from 3 days per week (Tier 2).

Fourth, small groups are recommended to maximize interactions with students including specific practice and corrective feedback. In this study, all groupings consisted of no more than three students; one group had only two students. In Table 1, we provide exemplars of how instruction was intensified in this study, which connects to Vaughn et al. (2012) recommendations.
Given that there is a group of young students who manifest severe mathematics difficulties, intensifying instructional variables is warranted to reduce poor performance on mathematics outcome measures and to increase mathematical understanding of early numeracy concepts and skills. Therefore, the intent of this study was to investigate the effect of a Tier 3 intensive intervention on early numeracy outcomes for second graders with severe mathematics difficulties. Three research questions guided this study:

1. What were the effects of an explicit, systematic Tier 3 (tertiary) intervention on timed progress monitoring measures of early numeracy mathematics (proximal measure) of students in the second grade with severe mathematics difficulties?

2. Did students with severe mathematics difficulties maintain their mathematics performance on timed progress monitoring measures of early numeracy mathematics (proximal measure)?

3. Did students receiving the early numeracy Tier 3 intervention demonstrate improved performance on a distal measure of mathematics concepts and skills?
Method

Participants

Thirty-three second graders participated in this study. Students who scored at or below the 10\textsuperscript{th} percentile on the district’s universal screener and benchmark test, the Texas Early Mathematics Inventories-Progress Monitoring (TEMI-PM) (Texas Education Agency/University of Texas System [TEA/UTS], 2008a), were selected to participate in the study contingent on signed informed parent consent and student assent as stipulated by Institutional Review Board research procedures. Second-grade teachers were responsible for administering the TEMI-PM, which occurred in winter before the study began and spring after the study concluded. For purposes of this study, we selected students who scored at or below the 10\textsuperscript{th} percentile on the screener (see Mazzocco, Feigenson, & Halberda, 2011) as participants. On the winter administration of the TEMI-PM, 48\% of the participants scored at the 1\textsuperscript{st} percentile, 12\% scored at the 6\textsuperscript{th} percentile, 9\% scored at the 3\textsuperscript{rd} percentile, 6\% were at the 5\textsuperscript{th}, 7\textsuperscript{th}, and 8\textsuperscript{th} percentiles, and 1\% scored at the 4\textsuperscript{th} and 10\textsuperscript{th} percentiles compared to their peers in second-grade who had an average winter raw score of 81.23 (27\textsuperscript{th} percentile) (TEMI-PM, Technical Manual, 2008b).

Twelve students were male, and 21 were female; one student was Caucasian, three were African American, and 29 were Hispanic. Thirty-two student received free or reduced meals. We report school and group demographic information in Table 2. Individual student data are available from the first author upon request.

Table 2. Demographic and Progress Monitoring Information for the Participants

<table>
<thead>
<tr>
<th>School/Group</th>
<th>N</th>
<th>Gender</th>
<th>Ethnicity</th>
<th>EL</th>
<th>F/RL</th>
<th>Baseline Mean (SD)</th>
<th>Intervention Starting/Ending Raw Scores</th>
<th>Intervention Mean (SD)</th>
<th>Maintenance Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1</td>
<td>3</td>
<td>F</td>
<td>1 B</td>
<td>3 No</td>
<td>2 Yes</td>
<td>56.92 (19.29)</td>
<td>64.7 to 83.7</td>
<td>73.3 (19.4)</td>
<td>98.0 (13.0)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>H</td>
<td></td>
<td>1 No</td>
<td>1 No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/2</td>
<td>3</td>
<td>F</td>
<td>2 M</td>
<td>1 B</td>
<td>3 Yes</td>
<td>65.80 (12.84)</td>
<td>70.0 to 101.0</td>
<td>85.1 (13.5)</td>
<td>95.0 (15.5)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>H</td>
<td></td>
<td>3 Yes</td>
<td>3 Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/1&amp;2</td>
<td>6</td>
<td>F</td>
<td>5 H</td>
<td>3 Yes</td>
<td>5 Yes</td>
<td>61.68 (16.47)</td>
<td>66.2 to 92.3</td>
<td>79.2 (17.6)</td>
<td>96.4 (17.4)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>M</td>
<td></td>
<td>5 No</td>
<td>1 No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/1</td>
<td>3</td>
<td>F</td>
<td>2 H</td>
<td>2 Yes</td>
<td>3 Yes</td>
<td>71.57 (9.21)</td>
<td>99.0 to 101.7</td>
<td>97.8 (14.6)</td>
<td>121.7 (13.0)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>No</td>
<td></td>
<td>2 Yes</td>
<td>1 No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/2</td>
<td>2</td>
<td>F</td>
<td>2 H</td>
<td>2 Yes</td>
<td>2 Yes</td>
<td>76.78 (9.85)</td>
<td>69.5 to 101.5</td>
<td>87.0 (10.4)</td>
<td>102.2 (7.0)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>F</td>
<td></td>
<td>1 No</td>
<td>2 Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/1 &amp; 2</td>
<td>5</td>
<td>F</td>
<td>5 H</td>
<td>4 Yes</td>
<td>5 Yes</td>
<td>73.61 (9.60)</td>
<td>79.3 to 101.6</td>
<td>93.2 (13.9)</td>
<td>113.9 (15.3)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>H</td>
<td></td>
<td>5 No</td>
<td>5 Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/1</td>
<td>2</td>
<td>F</td>
<td>1 B</td>
<td>1 Yes</td>
<td>2 Yes</td>
<td>70.75 (17.73)</td>
<td>79.0 to 93.0</td>
<td>86.4 (19.3)</td>
<td>100.5 (25.2)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>M</td>
<td></td>
<td>1 No</td>
<td>2 Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Setting

At the time of the study (according to state statistics), the school district had an enrollment of about 85,000 students with the following demographic characteristics: 51.5% male, 48.5% female; 11.3% African American, 3.6% Asian/Pacific Islander, 58.9% Hispanic, 0.3% Native American, and 25.8% White; 18.1% of the students participated in Bilingual Education, 19.1% Career and Technical Education, 9.9% English as a Second Language, 6.1% Gifted and Talented, and 9.5% Special Education. Also, 29% of the students were English learners, and 63.4% were designated as Economically Disadvantaged.

The participants attended second grade across five schools in a large urban school district in the Southwest. The schools were located geographically in the same quadrant of the school district. These schools shared similar demographic characteristics.

Students who met the criterion for inclusion in this study within each of the five schools from different or similar classes were as-
signed to small groups and taught by mathematics interventionists on a pull-out basis. Researchers assigned students to small groups based on (a) scheduling across general education teachers within and across schools, and (b) traveling time between schools for the research team mathematics interventionists. Small group Tier 3 intervention took place in empty classrooms or other spaces that would accommodate several students and the interventionist. There were two small groups in schools 1 and 2, three groups in school 3, and four groups in school 4; school 5 had only one group.

**Experimental Design**

A multiple baseline design across schools was employed (Kennedy, 2005). Performance on early numeracy mathematics skills and concepts measured by the Total Score of the Texas Early Mathematics Inventories-Aim Check (TEMI-AC; TEA/UTS, 2009), described in the Measures section, served as the dependent variable in this study. During the baseline phase, the TEMI-AC scores were collected to estimate trends and related patterns between groups within and across schools. For the two groups of students in School 1, mathematics interventionists administered five probes over a 5-school day period. Students in the two groups in School 2 also received five probes, but mathematics interventionists gave their probes across eight school days. Students in Groups 1, 2, and 3 in School 3 were administered six baseline probes over 11 school days. Mathematics interventionists gave eight probes across 14 school days to four groups in School 4; and for the fifth school with one group, ten probes were administered across 17 school days.

The intervention occurred for 54 days for the groups in Schools 1 and 2, and 59 days for the remaining schools and groups; the TEMI-AC was administered weekly during the intervention phase. A maintenance phase lasted two weeks for all groups in Schools 1, 2, 3, and one group in School 4; the remaining three groups in School 4 and the one group in School 5 had one week of maintenance due to the approaching end of the school year. The KeyMath-3 (Connolly, 2008), the distal measure, was administered before the baseline and after the maintenance phases to assess generalization.

**Measures**

**TEMI-AC.** The TEMI-AC (TEA/UTS, 2009) is a researcher-devised measure that contains four, 2-minute fluency subtests. Five alternate forms focus on Magnitude Comparisons, Number Sequences, Place Value, and Addition and Subtraction Combinations. The first subtest assesses Magnitude Comparisons (circle, from two numbers separated by a vertical dotted line, the number that is less or circle both numbers if they are equal). Magnitude Comparisons is similar in focus to Clarke’s and Shinn’s (2004) Number Identification Measure 1–20, but includes numbers from 0 through 999. Number Sequences, the second subtest, involves students writing the missing number from a three-number sequence; it is similar to Clarke’s and Shinn’s (2004) Missing Number Measure Blank Varied 1–20, but also includes numbers from 0 through 999. The third subtest, Place Value, involves students looking at pictorial depictions (base ten pictures) of hundreds, tens, and ones (ranging from 1 through 999) and writing the quantity depicted. The final subtest, Addition-Subtraction Combinations consisted of basic addition and related subtraction facts. The raw scores of the four subtests are summed to generate a total score that can be used to document progress. The TEMI-AC has five alternate forms; alternate forms reliability of the Total Score exceeds .80 across all forms. As a proximal measure, the TEMI-AC is a valid
measure to document student growth on early numeracy concepts and skills.

**KeyMath-3.** The *KeyMath-3* (Connelly, 2008) is a norm-referenced measure of mathematical concepts and skills that uses ten subtests to assess three broad content areas (Basic Concepts, Operations, and Applications); a total score is also available. Internal consistency (mostly in the .80s), alternate forms (from .74 to .92), and test-retest (coefficients were in the mid-.90s) were reported for the *KeyMath-3*. Test developers report considerable evidence of the validity of the *KeyMath-3* scores in the test’s technical manual.

**Procedures for the Intervention**

**Curriculum.** The purpose of the intensive intervention was to develop conceptual understanding (i.e., “comprehension of concepts, operations, and relations,” CCSS-M, 2010, p. 6) and procedural fluency (i.e., the ability to carry out procedures accurately and proficiently, CCSS-M, 2010) to improve competence with early numeracy concepts and operations measured by the *TEMi-AC*. The intensive intervention was adapted from a Tier 2 intervention (Bryant et al., 2008), which consisted of scripted lessons and progress monitoring measures. The curriculum focused on topics contained in the CCSS-M (2010) for kindergarten (counting and cardinality), and first and second grade (operations and algebraic thinking, and number and operations in base ten). Lessons emphasized number identification and writing, counting up to and back from 20, strategies for addition facts and related subtraction facts, magnitude comparisons (i.e., comparisons of numbers closer together on the number line and numbers further apart), number sequences (i.e., sequences that crossed decades and varied the location of the missing number [beginning, middle, end], and the base-10 system (i.e., place value). To reduce the “cognitive load,” the range of numbers 0 – 999 were chunked so that earlier lessons focused on 0 – 50, followed by 50 – 100, then groups of hundreds (e.g., 100 – 200, 200 – 300), and ending with the entire range mixed (i.e., 0 – 999). The lessons included these number ranges for number sequences (e.g., 199, ___, 201; ___, 329, 330), magnitude of numbers (which is less 419 or 491; 906 or 817), and place value (e.g., groupings of hundreds, tens and ones; composing and decomposing numbers). Addition and subtraction facts were taught using specific strategies (e.g., counting on [e.g., +2, +3; 6 + 2 = “6,” 7, 8 = 8], counting up to [e.g., 4 + ___ = 7], derived [e.g., 6 + 7 = 6 + [6 + 1] = [6 + 6] + 1 = 12 + 1 = 13], make ten plus more [8 + 4 = 8 + [2 + 2] = [8 + 2] + 2 = 10 + 2 = 12] to build fluency; problems were presented vertically and horizontally. The concept of equality and the commutative property were stressed as students solved equations.

**Tier 3 intervention.** The lessons were designed to promote the development of conceptual and procedural knowledge through multiple mathematical representations (e.g., connecting cubes, base ten models, number lines, ten frames, diagrams, symbols) and a combined approach of strategy instruction and explicit, systematic instruction; which consisted of modeling, providing multiple practice opportunities, scaffolding instruction with sequenced examples and prompts and cues, checking for understanding after guided practice (i.e., “Teach me how to do this problem”), and utilizing error correction procedures (Coyne et al., 2011; Gersten, Chard, et al., 2009).

The instructional routine was composed of five instructional components in each lesson including *Warm-Up* (i.e., activating background knowledge [counting up to and back from 20, counting by 10s and 5s] and reviewing previously taught addition facts and related subtraction facts [distributive practice – reviews that take place over
time], Preview (providing an advance organizer for the lesson – “Today we will focus on how to use strategies we have learned to solve addition and subtraction facts.”), Modeled Practice (teaching the concepts and skills with student interactions, using “think aloud” to make the process visible to students), Guided Practice (practicing as a group [choral and individual responding] with the interventionist), and Daily Check (independent practice for progress monitoring purposes; see Bryant et al., 2011), which assessed the content taught each day for each lesson. Before the Daily Check occurred, the teacher provided a review statement (e.g., “Remember to use your strategies not your fingers.”) so that students could recall the lesson’s content before taking the Daily Check. Then the mathematics interventionists gave students 2 minutes to complete the items independently; students checked their work with the teacher and corrected errors.

Trained mathematics interventionists taught two lessons for 30 min a day Monday – Thursday for eight weeks; maintenance checks and generalization testing were conducted four weeks at the conclusion of instruction. Fridays were devoted to a game format, which incorporated game card items related to the content taught. Research findings have shown that a game-like format (i.e., board game) helps improve early numeracy concepts and skills of low performing students (Siegler, 2009). Games that consisted of concepts and skills from the intervention, lasted 20 to 25 min. Then, to help motivate students, based on the number of accurate answers to questions on the cards, they were awarded prizes. The mathematics interventionists administered a 2-min progress monitoring measure (TEMI-AC) following the game. The grouping structure was two or three students per interventionist. The interventionists had tutored in previous early numeracy projects with struggling students, and thus, were well trained on the curriculum, delivery, and progress monitoring.

**Tier 1 Core Instructional Content and Procedures**

All participants received Tier 1 core instruction in their general education classrooms. In response to a self-report online survey, the students’ teachers indicated they often adapted core instruction for their struggling students. Adaptations included having students use concrete manipulatives, and pairing the struggling students with peers for assistance. Respondents also noted that they provided their struggling students with more instruction and time to respond, as compared to their typically achieving peers. Teachers provided individual remedial instruction two to three times during the week. Despite this differentiation and extra practice all of the participants scored at or below the 10th percentile on the winter administration of the district’s school-wide assessment measure (TEMI-PM). Clearly, at the mid-point of the school year, the participants in this study were not adequately responding to differentiated instruction as part of core mathematics instruction and remained as low responder (Vaughn et al., 2009).

**Fidelity of Intervention Implementation**

Research team members conducted fifteen observations using the Fidelity of Implementation (FOI) form, which was a researcher-created fidelity observation scale. Two members of the project staff conducted the observations. We trained each member in using the fidelity scale to conduct an observation. Inter-rater reliability was 80% agreement as computed by the number of agreed upon fidelity items ÷ total number of fidelity items; differences were discussed and resolved. Each observer rated the extent to which each lesson component was implemented with respect to
different aspects of the intervention including (a) reviewing the basic facts, (b) reviewing prerequisite skills for the lesson, (c) providing modeled practice, (d) providing guided practice, and (e) having students complete the independent practice. Team members observed two lessons during each fidelity check; the observer also recorded comments, which the observer shared with the mathematics interventionists. The range of scores was a low of 23 to a high of 37 (median = 30). The scores were then used to assign an overall Fidelity Rating on a 1-to 5-point scale, where 1 = Poor and 5 = excellent. Overall average ratings ranged from a low of 2.5 to a high of 4.5; the median FOI score across the observations was 3.43, indicating that the mathematics interventionists administered the lessons with good fidelity.

Results

In single-case research, results are interpreted using visual inspection of the data. According to Kratochwill et al. (2010), visual inspection of the data should involve an examination of commonly accepted practices including an analysis of level, trend, variability, immediacy of effect, and overlap. We provide results for the visual analyses (see Figure 1), experimental control, effect sizes, and generalization.

Figure 1. TEMI-AC Weekly Data for Five Schools
Visual Analysis

The total scores averaged for each group within each school on the TEMI-AC across the baseline, intervention, and maintenance conditions are shown in Figure 1. Group 1 of School 1 (Figure 1, top panel; closed circles) demonstrated scores at relatively low and stable levels, with a slight upward trend initially, throughout the baseline condition ($M = 55.8$; range of baseline data points = 44 to 60.7). Group 1’s scores continued at similar levels initially upon the implementation of the intervention before increasing on an upward trend throughout the remainder of the intervention condition ($M = 73$; range 60.5 to 83.7). Group 1’s scores remained high following the removal of the intervention during maintenance ($M = 99.8$; range = 91 to 108.5). Group 2 of School 1 (Figure 1, top panel; open circles) demonstrated scores that were relatively low during baseline and, following a slight increase in level, continued at relatively low and stable levels throughout the baseline condition ($M = 63.5$; range, 55.3 to 69.3). Group 2’s scores increased somewhat during the first session of the intervention condition ($M = 85.1$; range 70 to 101) and increased further during subsequent sessions. Scores continued to occur at relatively stable levels above those observed during baseline throughout the remainder of the intervention condition. Group 2’s scores remained relatively high during maintenance ($M = 95$; range = 94.7 to 95.3).

Group 1 of School 2 (Figure 1, second panel; closed circles) demonstrated relatively low and stable scores throughout the baseline condition ($M = 71.1$; range, 65.3 to 74.7). Group 1’s scores immediately increased upon the implementation of the intervention ($M = 97.8$; range 83 to 114.3) and, following a slight decrease, continued at consistent and increasing levels above those observed during baseline throughout the remainder of the intervention condition.

Group 1’s scores increased to higher levels during maintenance ($M = 121.7$; range = 120 to 123.3). Group 2 of School 2 (Figure 1, second panel; open circles) demonstrated relatively low but increasing scores during the baseline condition ($M = 75.6$; range, 65 to 86) with a decrease observed during the final session of baseline. Group 2’s scores initially decreased upon the implementation of the intervention condition ($M = 87.4$; range 69.5 to 101.5) but increased throughout the remainder of the condition to relatively stable levels that were higher than those observed during baseline. Group 2’s scores continued at similar levels during maintenance ($M = 102.3$; range = 98.5 to 106).

Group 1 of School 3 (Figure 1, third panel; closed circles) demonstrated relatively low and stable scores throughout the baseline condition ($M = 70.8$; range, 63.5 to 77.5). Group 1’s scores increased slightly upon the implementation of the intervention condition ($M = 86.4$; range 75 to 97.3) and continued at similar and stable levels for the subsequent four sessions before increasing to levels above baseline throughout the remainder of the intervention condition. Group 1’s scores increased to slightly higher levels during maintenance ($M = 100.5$; range = 97.5 to 103.5). Group 2 of School 3 (Figure 1, third panel; open circles) demonstrated relatively low and consistent scores during the baseline condition ($M = 68.4$; range, 64.3 to 73). Group 2’s scores increased slightly upon the implementation of the intervention condition ($M = 85.3$; range 74.3 to 101.3) to levels above baseline. Group 2’s scores continued to increase during the remainder of the intervention condition. Group 2’s scores increased to slightly higher levels during maintenance ($M = 104.9$; range = 101 to 108.7). Group 3 of School 3 (Figure 1, third panel; closed triangles) demonstrated relatively low but slightly increasing scores initially during the baseline condition before occurring at relatively stable
levels during the final three sessions of the condition ($M = 63.4$; range, 52 to 72). Group 3’s scores initially decreased upon the implementation of the intervention condition ($M = 89.3$; range 64.7 to 109.3) but increased throughout the remainder of the condition to levels that were considerably higher than those observed during baseline. Group 3’s scores continued at similar levels during maintenance ($M = 110.4$; range = 108 to 112.7).

Group 1 of School 4 (Figure 1, fourth panel; closed circles) initially demonstrated relatively low and stable scores and eventually an increasing trend during the baseline condition ($M = 60$; range, 54 to 70.3). Group 1’s scores increased slightly upon the implementation of the intervention ($M = 83.6$; range 72.7 to 98.7) and continued to increase in a steady trend at levels above baseline throughout the remainder of the intervention condition. Group 1’s score decreased slightly during maintenance (91.7). Group 2 of School 4 (Figure 1, fourth panel; open circles) demonstrated stable scores in a slightly increasing trend during the baseline condition ($M = 91.6$; range, 84.5 to 98). Group 2’s scores increased slightly upon the implementation of the intervention ($M = 101.5$; range 92 to 112.5), decreased to levels below those observed during baseline for three sessions (sessions 12-14), and then increased during the remainder of the condition to levels above those observed during baseline. Group 2’s score increased considerably during maintenance (126). Group 3 of School 4 (Figure 1, fourth panel; closed triangles) demonstrated initially low scores that increased across the duration of the baseline condition ($M = 77.3$; range, 61.3 to 87). Group 3’s scores increased slightly upon the implementation of the intervention ($M = 103.6$; range 87.7 to 117) and, following a subsequent decrease during one session (session 10), continued to increase consistently above baseline levels during the remainder of the condition. Group 3’s score increased considerably during maintenance (123). Group 4 of School 4 (Figure 1, fourth panel; open triangles) demonstrated relatively stable scores with a flat trend during the baseline condition ($M = 85.1$; range, 67 to 91.3). Group 4’s scores immediately increased upon the implementation of the intervention ($M = 108.3$; range 100 to 122) and continued to increase consistently above baseline levels during the remainder of the condition. Group 4’s scores continued at similar levels during maintenance ($M = 121$; range 121.7 to 120.3).

Group 1 of School 5 (Figure 1, fifth panel; closed circles) demonstrated initially low scores that increased to stable levels during the baseline condition ($M = 92.6$; range, 77 to 104). Group 1’s scores continued at similar levels upon the implementation of the intervention ($M = 115.4$; range 101 to 131) and, following a stable and flat trend during the first five sessions of the condition, increased to levels above baseline during the remainder of the condition. Group 1’s score was at a similar level during maintenance (126).

**Experimental Control**

Experimental control for multiple baseline designs involves the staggered introduction of a treatment across participants (or groups of participants) and the staggered demonstration of positive effects of the treatment. The demonstration of experimental control is evident when positive changes are observed only following the introduction of the intervention, and there is evidence of consistently low levels of responding during baseline conditions across three or more data series (see Horner et al., 2005). Figure 1 shows a design that consists of twelve data series across five schools and the implementation of the intervention at a different point in time for each school. For the majority of the groups within and across
schools, students’ response patterns show increases in scores during the intervention.

Thus, the majority of the results show experimental control by demonstrating a co-variation between change in behavior patterns and introduction of the intervention across five schools at five different points in time. According to Kratochwill et al. (2010), a minimum of three replications across different cases of an experimental effect “meets evidence” for single-case design standards.

**Effect Sizes**

The Nonoverlap of All Pairs (NAP; Parker & Vannest, 2009) procedure was used to compute effect sizes. NAP is an extension of Mastropieri’s and Scrugg’s (1985-1986) Percentage of Nonoverlapping Data (PND), except instead of comparing each intervention data point with the highest baseline point, as is the case with PND, NAP compares all data points during the intervention to all baseline data points for overlap to provide a valid effect size. NAP, like PND and other effect size calculations, are meant to supplement, not supplant, visual inspection. Having already examined how visual analysis supports the validity of the intensive intervention, we also report NAP data for all groups and schools in the study (see Table 3). We note overlap of data in three groups (Group 2, School 2; Group 3, School 3; Group 1, School 5); no overlapping data are evident in the remaining 9 groups.

**Table 3. Results of Effect Size Analysis Using Percentage of Nonoverlapping of All Pairs (NAP)**

<table>
<thead>
<tr>
<th>School/Group</th>
<th>NAPa</th>
<th>School/Group</th>
<th>NAP</th>
<th>School/Group</th>
<th>NAP</th>
<th>School/Group</th>
<th>NAP</th>
<th>School/Group</th>
<th>NAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1</td>
<td>94.5%</td>
<td>2/1</td>
<td>100%</td>
<td>3/1</td>
<td>98.3%</td>
<td>4/1</td>
<td>100%</td>
<td>5/1</td>
<td>97.5%</td>
</tr>
<tr>
<td>1/2</td>
<td>100%</td>
<td>2/2</td>
<td>86%</td>
<td>3/2</td>
<td>100%</td>
<td>4/2</td>
<td>82.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/1 &amp; 2</td>
<td>94.5%</td>
<td>2/1 &amp; 2</td>
<td>93%</td>
<td>3/3</td>
<td>94.1%</td>
<td>4/3</td>
<td>85%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/1 &amp; 2</td>
<td>98.3%</td>
<td>4/4</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp; 3 &amp; 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aNonoverlapping of All Pairs

**Generalization**

The *KeyMath-3* (Connolly, 2008), the distal, generalization measure, was administered before the intervention occurred and following the maintenance condition. Of the 33 students who completed the intervention and were administered the *KeyMath-3*, the average total standard score at pretest was 74.41, with standard scores ranging from a low of 59 to a high of 88; five students had a post-test score at or above the 25th percentile. One student in School 2, Group 1 had a pretest standard
score of 80 and a posttest score of 95. One student in School 3, Group 1 had a pretest of 84 and scored a 90 at posttest. In School 4, Group 1, one student achieved a 77 at pretest and a 95 at posttest. Two students in School 5, Group 1 achieved posttest standard scores at or below 90; one student’s pretest was a standard score of 83 with a posttest of 91, and the other student scored a 79 at pretest and 92 at posttest. Although the number of students who achieved a standard score of 90 or higher is small, it is worth noting that all of these students scored at or below the 21st percentile on the initial KeyMath-3 administration and had scored at or below the 10th percentile on the winter administration of the district’s benchmark measure (TEMI-PM). We provide pretest and posttest standard scores in Table 4.

**Table 4. Pretest and Posttest Results for the KeyMath-3**

<table>
<thead>
<tr>
<th></th>
<th>Mean Form A</th>
<th>Mean Form B</th>
<th>Standard Error</th>
<th>t-score</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Concepts</td>
<td>74.59</td>
<td>81.17</td>
<td>1.179</td>
<td>5.585</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>Operations</td>
<td>79.55</td>
<td>84.79</td>
<td>1.417</td>
<td>3.698</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>Applications</td>
<td>75.97</td>
<td>84.52</td>
<td>1.451</td>
<td>5.894</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>Total Score</td>
<td>74.41</td>
<td>81.48</td>
<td>1.094</td>
<td>6.461</td>
<td>p &lt; .01</td>
</tr>
</tbody>
</table>

*a Standard Scores*

**Discussion**

The intent of this study was to investigate the effect of a Tier 3 intensive intervention on early numeracy outcomes for second graders with severe mathematics difficulties. The researchers employed a multiple baseline single-subject design, with twelve groups of students in five schools. Research findings show that mathematics outcomes improve when students are provided explicit, systematic instruction on critical early numeracy concepts and skills that includes strategies, systematic review, and multiple opportunities for practice with ongoing, regular progress monitoring (L. S. Fuchs, D. Fuchs, & Compton, 2012; L. S. Fuchs et al., 2010; Hunt, Valentine, Bryant, Pfannenstiel, & Bryant, 2016).

**Effects of the Tier 3 Intervention**

Overall, findings from the visual analyses are positive yet somewhat tempered because in two cases (groups) (School 4, Groups 1 and 3), we observed upward trends during baseline, which continued when the mathematics interventionist taught the intervention. However, as noted in the Council for Exceptional Children’s Standards for Evidence-Based Practices in Special Education (2014), “A single-subject study is considered to have positive effects when a functional relationship is established between the independent and dependent variables, resulting in a meaningful, therapeutic change in the targeted dependent variable for at least three-fourths (75%) of the cases” (p. 7). Thus, the findings of this study demonstrate that 75% of the cases met this mini-
In looking at the patterns of responding during the intervention across the groups, a slow, upward trend for some groups is noted, suggesting that for low responders such as Tier 3 identified students, progress at a lower rate or trend than higher responders is not unusual when an intervention is first introduced (Vaughn et al., 2009). This slower rate of progress should not be surprising given that for some low responders, improvement may take time in an intensive intervention condition and be gradual given some students’ patterns of chronic, weak, “resistant” performance. For example, in examining the trend, Group 1, School 1 had the lowest starting baseline data point and exhibited slow progress until the conclusion of the study. By the end of the study, the groups’ ending intervention scores on the TEMI-AC ranged from the 16th (Group 1, School 1) – 78th (Group 1, School 5) percentile with a mean percentile of 41.5 across groups. All of the groups except one (Group 1, School 1) showed intervention ending scores in the average range; the one exception scored below average.

Researchers have identified risk factors for early, severe mathematics difficulties including low mathematics performance at the end of kindergarten, and socioeconomic status of families (Jordan, Kaplan, Ramine ni, & Locuniak, 2009). For example, Morgan, Farkas, and Wu’s (2009) longitudinal work on mathematics trajectories of students with persistent mathematics difficulties showed that severe mathematics deficits over several years would likely continue into the elementary grades. Encouragingly, in an earlier study, B. Bryant et al. (2014) designed an intensive intervention for a trained mathematics interventionist to conduct in three elementary schools located in a school district in Central Texas. Twelve second-grade students who had received two rounds of Tier 2 intervention in first grade were re-evaluated through the district’s universal screener and scored at or below the 10th percentile, which “is indicative of chronically severe unresponsive performance.” (p. 3). Much like the current study, the students received an intensive Tier 3 intervention that was characterized as consisting of small teacher-student ratios, 30-min sessions with two lessons four days per week for 10 weeks. Also, the mathematics interventionist played a game with the students on skills from the intervention, and conducted progress monitoring after the game. The intervention in the B. Bryant (2014) and the current Tier 3 study consisted of the same components and the same measures (see B. Bryant 2014 for details). Findings from the B. Bryant et al. study showed “with regard to factors of visual analysis, with all three groups, low and stable levels of responding were observed during baseline, an immediate intervention effect was observed, and stable and increasing trends in responding were observed during the intervention” (p. 7). Whereas, the findings from the current study were positive but not as strong as the findings from the B. Bryant et al. (2014) study. However, students in the current study manifested identified mathematics difficulties in second grade; thus, we can hypothesize that these difficulties were likely evident in earlier grades due to the persistent nature of their poor mathematics performance and because they had only received differentiated rather than intensive interventions during the fall semester of second grade. Overall, though, results from this current study are encouraging considering that younger students with chronic mathematics difficulties typically are unlikely to catch up to their typically achieving peers and, thus, remain at a decided disadvantage learning new mathematics skills and concepts, which are based on a solid early numeracy foundation (Jordan et al., 2009; Morgan et al., 2009).
Another risk factor is low socio-economic status. Children from low-income families often come to school without skills that prepare them for formal school mathematics instruction (Clarke et al., 2016; Dyson et al., 2013; Siegler, 2009). In their recent work, Morgan, Farkas, Hillemeier, and Maczuga (2016), found that “low family SES is an overwhelmingly important predictor of persistent mathematics difficulties during the preschool, elementary, and middle school grades” (p. 315). Thus, the need for researchers to address the importance of strong core instruction (e.g., Clarke et al., 2011) and early consistent, intensified intervention for students from low SES families beginning in preschool and kindergarten and continuing throughout the elementary grades cannot be overestimated (Dyson et al., 2013). Given that the majority of the students in this study were eligible for free/reduced meals, it is heartening to see growth in mathematics performance. It is possible that intensified intervention components (e.g., group size, multiple representations, memory enhancements, games) such as those in this study can make a difference in mathematics outcomes for many students with initially low mathematics performance to help them learn early numeracy concepts and skills (i.e., magnitude of numbers, sequencing of numbers, place value, and basic addition and subtraction facts) (Vaughn et al., 2012).

Effect sizes were calculated to determine the magnitude of improvement; we did not observe evidence of overlapping data points between baseline and intervention for nine out of twelve groups. According to Parker and Vannest (2009), interpretation of NAP results includes weak effects (0 – 65%), medium effects (66 – 92 %), and strong effects (93 – 100%). All results show NAPs in excess of 80%, with 9 of 12 percentages exceeding .90. At the school level (all groups within each school), all five NAPs exceed 90%; when combined with the visual inspection of the data, NAP statistics provide additional support for the evidence of the intensive interventions for most of the student groups. This finding is in line with previous studies that showed positive effect sizes for intensifying mathematics interventions (Hunt et al., 2016; Clarke et al., 2016).

**Maintenance Effects**

Results from the TEMI-AC maintenance data showed 10 out of 12 groups with performance either at or above the intervention scores. These findings are encouraging given that all of the students scored at or below the 10th percentile on the winter administration of the school district’s benchmark measure. Students who had received a semester (fall to winter) of differentiated core instruction in the winter continued to demonstrate significantly depressed mathematics performance; this type of low responding is predictive of continued mathematics poor performance without an intensified intervention (Jordan et al., 2009). Given the low percentile ranking before the intervention, it is reassuring to note that all of the groups’ maintenance mean scores showed the ability to maintain growth.

**Generalization Effects**

Findings on the KeyMath-3 (Connolly, 2008; generalization measure) indicated that five out of the 33 students were performing in the average range of mathematics achievement at the conclusion of the study. For students with severe mathematics difficulties, the ability to generalize their learning to broad-based measures will likely take time and require a longer duration of intervention than what was possible in this study. It is encouraging that at least a small group of students did perform in the average range on the KeyMath-3 suggesting that some students with severe difficulties
can generalize learning from an early numeracy intensive intervention to a broad-based measure.

In the current study, we targeted the development of whole number concepts. The mathematics intervention did not include the other mathematics domains measured on the KeyMath-3 (e.g., Applications). In a previous randomized controlled trials study (Bryant et al., 2011) of an early numeracy Tier 2 intervention, results on the Stanford Achievement Test-Tenth Edition ([SAT-10], Harcourt, 2003) Mathematics Procedures measure were not statistically significantly different between the treatment and comparison conditions. However, Clarke et al. (2016) found in their efficacy study of a Tier 2 kindergarten mathematics intervention, statistically significant differences favoring the treatment group on the distal (generalization) measure, the Test of Early Mathematics Abilities-Third Edition ([TEMA-3], Ginsburg & Baroody, 2003), but not on the Early Numeracy-Curriculum Based Measure ([EN-CBM], Clarke & Shinn, 2004); thus, showing mixed results for generalization to a broad base or distal measure. Therefore, empirical research is needed to determine ways to promote the occurrence of generalization of mathematics knowledge on skill and concepts from proximal to distal measures and whether there are groups of students performing at various percentile rankings whose learning does and does not generalize across types of measures.

Limitations and Future Research

Several limitations are noted for this study and contribute to the tempering of the overall findings. First, as previously noted, two groups’ baseline scores showed an upward trend, even though one of them might be attributed to a missing score in the second to last baseline test. Thus, we did not observe experimental control with these two groups. Second, in examining the fidelity of implementation scores, one mathematics interventionist’s average score was 2.5. Despite regular meetings with the interventionists to resolve implementation issues, fidelity remained low for one interventionist. Finally, being mindful that each group included two to three students and data points represented the average of the TEMI-AC scores, it is possible that a subgroup of the total sample required another round of Tier 3 intervention or referral for special education testing. Separating each group’s scores to show individual student scores would likely shed light on the possibility of a subgroup of students whose performance requires further diagnosis and intensity of intervention; however, this is beyond the scope of this study.

Regarding future research, first, more studies are needed to determine ways to enhance generalization of findings from proximal to distal measures. Obtaining improved mathematics performance on proximal measures is a good first step. Ways to promote generalization of findings for students with severe mathematics difficulties to distal, or broad-based measures is crucial to understand and could influence how researchers design interventions to foster the occurrence of generalization. Second, large-scale studies employing the intervention from this study are necessary to determine how larger samples of similarly performing students with this study improve their mathematics performance. Third, studies involving students with severe mathematics difficulties from low SES families are imperative. Given Morgan et al. (2016) longitudinal results, this research focus is critically important. Finally, a large portion of the sample were English learners. Even though the students were receiving their core mathematics instruction in English, a requirement for the current study, it is difficult to say what the effects of their language origin might have had on their per-
formance on the intervention. Future research that examines these effects would be useful.

Implications for Educational Practice

We provide several implications for practice. First, students must be screened and identified at an early age for the presence of mathematics difficulties and an intensive intervention should be employed beyond differentiating core instruction. Student progress should be monitored frequently beyond the typical beginning, middle, and end of year screening and benchmark assessments.

Second, decision-makers need to consider whether students with severe mathematics difficulties (i.e., performing below the 10th percentile) should receive Tier 3 intervention rather than beginning in Tier 2. This “wait to fail” model further exacerbates persistently low mathematics performance that interventionists could target earlier. School districts will likely need to decide how to distinguish those students who are good candidates for Tier 3 so that they can receive a more intensified intervention upon identification.

In sum, overall findings show promise for the Tier 3 intensive intervention implemented in this study. Future research can play an important role in further identifying how interventions can be intensified and ways to promote the occurrence of generalization to broad-based mathematics measures.

References


intensity. *TEACHING Exceptional Children, 50*(1), 35-43.


http://www2.ed.gov/about/bdscomm/list/mathpanel/index.html


Texas Education Agency/University of Texas System, (2008a). Texas Early Mathematics Inventories - Progress Monitoring. Austin, TX: Authors


Texas Education Agency/University of Texas System, (2009). Texas Early Mathematics Inventories – AIM Check. Austin, TX: Authors

