The Dissertation Committee for Sarah Adair Benz Certifies that this is the approved version of the following dissertation:

The Effect of Behavior Within a Word-Problem Intervention for Students with Mathematical Difficulties

Committee:

Sarah R. Powell, Supervisor

Diane P. Bryant

Christian Doabler

Candace Mulcahy

The Effect of Behavior Within a Word-Problem Intervention for Students with Mathematical Difficulties

by

Sarah Adair Benz

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2018

Acknowledgements

I would like to acknowledge all those who have supported me through my Ph.D. First and foremost, it has been a privilege to work with my advisor and dissertation committee supervisor, Dr. Sarah Powell. She has provided me with support, guidance, and mentoring beyond my expectations. I would like to acknowledge the other members of my dissertation committee who thoughtfully spent their time and efforts to provide valuable feedback throughout this process. I also would like to acknowledge Dr. Kate Berry for all of her hard work and encouragement throughout my dissertation. To my parents for always inspiring me to be the best that I can be, no matter the field and for supporting me through the countless moves in pursuit of my dreams. To my brothers, for their unwavering love, support, and laughter. Lastly, to Nicholas for the love, support, understanding, and sacrifices that you have made as I pursued my goals.

The Effect of Behavior Within a Word-Problem Intervention for Students with Mathematical Difficulties

Sarah Adair Benz, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Sarah R. Powell

Students with or at-risk for challenging behaviors (e.g., students with emotional and behavioral disorders [EBD]) represent a group of children who often experience difficulty with both behavior and academics. Researchers and educators have designed interventions to improve the conduct for students with challenging behaviors, however, few mathematics interventions are specifically designed to support students with challenging behaviors. One mathematics intervention with a strong evidence base for use with students with learning disabilities is using schemas to solve word problems. Notably, researchers also have identified schema-based instruction as an effective intervention to solve word-problems for students with EBD. While a promising intervention, little is known about the effectiveness of a word-problem intervention relative to student behavior. That is, do behavioral profiles collected before the start of intervention predict response to intervention? This study sought to compare the behavioral profiles of students with and without mathematics difficulty (MD) and investigate the efficacy of a randomized control trial of a word-problem intervention for students with MD. Results of the present study indicated that students with MD demonstrate higher externalizing and internalizing behaviors than students without MD. Additionally, students in the word-problem intervention groups outperformed their peers in a no-treatment comparison group on all outcome measures, but most notably, on double-digit word problems with an effect size of 1.01. While students with high externalizing scores in the treatment condition performed significantly lower on a double-digit word-problem outcome measure than students without high externalizing scores, their externalizing score was not necessarily predictive of their word-problem gains. When comparing students with and without high internalizing scores in the treatment group, there was no statistical difference. Findings from this study establish a need to continue to study the link between mathematics difficulties and behavioral challenges.

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

The National Assessment of Educational Progress (NAEP, 2015) reported that only 40% of fourth-grade students and 33% of eighth-grade students in the United States met proficiency levels in mathematics. More alarming, students with disabilities performed significantly lower, with only 16% and 8% of fourth- and eighth-grade students with disabilities, respectively, scoring proficient or above proficient. Students with disabilities constitute approximately 13% of the total school-aged population (U.S. Department of Education, 2016) and therefore, require significant and specific attention to support mathematics proficiency.

Over the last 25 years, researchers have identified programs and interventions that are considered evidenced-based practices for students with or at-risk for mathematics difficulties (MD). For example, Peer Assisted Learning Strategies (PALS) Math (Fuchs et al., 1997; Fuchs et al., 2002) and Early Learning in Mathematics (ELM; Chard et al., 2008) are just two of the programs that have been established as interventions that support mathematics learning through explicit teaching and practice of concepts and procedures. While these interventions prove effective for specific students with disabilities (i.e., students with MD), a significant need exists to develop interventions to better support students with disabilities who continue to struggle in mathematics.

Students with or at risk for emotional and behavioral disorders (EBD) represents a subgroup of students who experience difficulty in mathematics, despite the use of evidence-based practices (Mulcahy & Krezmien, 2009). Not only do students with or at-risk for EBD struggle academically similar to students with MD, they also exhibit unique and challenging behaviors. These challenging behaviors may hinder students' abilities to access and benefit from established mathematical interventions might be unsuccessful (Lane, Barton-Arwood, Nelson, & Wehby, 2008).

The Individuals with Disabilities Education Act (IDEA, 2004) requires all students with disabilities to be educated in their least restricted environment (LRE) including students with EBD. Presently, more than 80% of students with EBD receive at least *some* of their core instruction in the general education classroom (U.S. Department of Education, 2016). Little is known, however, about the impact of student behavior on the efficacy of mathematical interventions. Typically, mathematics interventions are validated using group-design research methods. Despite the methodological rigor provided by those methods, group design studies often lack generalizability to specific student populations (Hedges, 2013). Maggin, Wehby, Farmer, and Brooks (2016) proposed that interventions from experimental group-based research may benefit from a process of adaption and individualization to improve effectiveness for particular students. Students with co-occurring academic and behavioral issues may require additional and individualized support over and above typical evidence-based practices.

Statement of Purpose

The purpose of the current study was to examine the behavioral profiles of students with and without mathematics difficulty (MD) and investigate the efficacy of a randomized control trial of a word-problem intervention for students with MD. First, this study compared the behavioral profiles of students with and without MD. Second, this study investigated the efficacy of a schema-based intervention (e.g., Pirate Math) for improving word-problem solving performance for third-grade students with MD. That is, in general, do students that receive the word-problem intervention outperform control students who did not receive the word-problem intervention? Last, this study attempted to determine if response to a word-problem intervention (e.g., Pirate Math) was predicted by student behavioral profiles. That is, did students that exhibit high rates of at-risk behaviors respond differently to the word-problem intervention versus students that did not display high rates of at-risk behaviors?

Pirate Math is an evidence-based intervention that is developed for elementary-aged students who experience difficulty in word-problem solving. Little is known, however, about the efficacy of Pirate Math for third-grade students that exhibit high rates of at-risk behaviors (i.e., externalizing behaviors or internalizing behaviors). With this study, I contribute to the growing body of research that examines the effects of randomized-control trial interventions on a particular subgroup of students (i.e., students with or at-risk for challenging behaviors comorbid with MD).

Chapter II: Review of the Literature

This review of the literature summarizes key research in the fields of mathematics and behavior. First, I examine the research on students with MD, particularly the prevalence and characteristics of students with MD, along with mathematics interventions that may be helpful for students with MD, with a focus on word-problem interventions. Then, I discuss research about students with challenging behaviors, the difficulty those students face with mathematics, and mathematics intervention research for students with challenging behaviors. Finally, I describe the theoretical significance of the mathematical intervention used in this study (i.e., Pirate Math).

Students with Mathematics Difficulties

While not all students with disabilities have difficulty with mathematics, many of them do. Students with disabilities, however, display uneven patterns of MD. The disability category with the lowest rates of MD are students with speech impairments (22%), while students with intellectual disabilities display the highest rates of MD (84%; Blackorby et al., 2004). In this paper, we focus on students who have persistent difficulty in mathematics. In the literature, these students may be referred to as having a mathematics disability or dyscalculia. I, however, will intentionally refer to this diverse group of learners as students with or at-risk for MD. This term may encompass students that have mathematics disabilities, dyscalculia, or mathematics difficulties.

Students with MD are typically considered students to be low performing in mathematics or have difficulty learning mathematics (Geary, 2011). MD can be identified in several ways. Students can score in a lower-than-average range (e.g., between the 10th and 25th percentile) or below a cut score (e.g., below the 15th percentile) on mathematics achievement tests (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007). Students may also demonstrate mathematics performance significantly below average (e.g., 2 SD below the mean). Students can also be identified as experiencing MD through a response to intervention process in which students demonstrate persistent low performance and limited response to high-quality instruction (Geary, 2011).

Studies on the potential etiology of MD reveal that genetic and environmental conditions contribute to students' mathematics achievement, and MD is often comorbid with reading disability or attention deficit hyperactivity disorder (ADHD; Barbaresi et al., 2005). Students with MD may experience negative post-school outcomes as well, noting lower rates of full-time employment and more frequent periods of unemployment (Geary, 2011).

Prevalence. Estimates of MD vary significantly. In terms of mathematics difficulty, approximately 25% to 35% of students persistently struggle with mathematics (Mazzocco, 2007). An additional 5% to 8% of school-aged students have a mathematics disability and experience significant deficits that require specially designed instruction (Geary, 2004). Students can encounter difficultly in any area of mathematics, including counting knowledge, mathematics fact retrieval, or word-problem solving (Geary, Hoard, Bailey, 2012). In one study, researchers examined more than 300,000 students in 5th grade between 1987 and 2007 and identified 5% to 7% of those students were between three to four grade levels behind in mathematics (Barbaresi et al., 2005). With this knowledge, it is necessary to provide additional mathematics support to students with MD.

Students with Challenging Behaviors

Another group of students who may experience difficulty in school are students with challenging behaviors. Students that present with challenging behaviors, such as students with EBD or ADHD, are included in the heterogeneous group of students with MD. The following includes a summary of research on students that exhibit challenging behaviors.

Prevalence. Nearly all school students experience circumstances or contexts where socially or emotionally students struggle to regulate their behavior (Maggin et al., 2016). Studies suggest that more than 30% of school-aged students will experience a mental health difficulty during their educational career (i.e., anxiety, depression, or trauma; Forness, Freeman, Paparella, Kauffman, & Walker, 2012). Many students experience momentary emotional and behavioral problems, but with proper support, adequately adjust and experience positive post-school outcomes (Forness et al., 2012). However, approximately 5% to 13% of the school-age population will experience sustained emotional or behavioral difficulties that have long-term implications (Maggin et al., 2016). Students with an EBD do not automatically qualify for services under the Emotional Disturbance category in special education (Forness et al, 2012). In order to qualify for ED, the student must meet the federal definition outlined in the regulations. The federal regulations limit ED disability identification to only students whose EBD would affect their academic achievement. In this dissertation, therefore, the term *challenging behaviors* to describe any student with an emotional, behavioral, or psychiatric disorder listed in the American Psychiatric Association (2013) diagnostic manual. Disorders may include, for example, students with conduct disorders, attention-deficit/hyperactivity disorder (ADHD), depression, and anxiety. The term *challenging behaviors* may include students with a diagnosed disability for ED or students that are displaying at-risk behaviors that may worsen without the support of interventions.

Characteristics. Students with challenging behaviors are characterized by behaviors that negatively influence their ability to succeed academically (Lane et al., 2008). Typically, students with challenging behaviors, especially EBD, earn lower grades and have higher rates of school dropout versus other students with disabilities (Wagner et al., 2004). Further, students with challenging behaviors experience increasingly negative post-school outcomes, difficulty with problem solving, and engage in high rates of substance abuse (Bullis & Yovanoff, 2006).

Students with challenging behaviors may exhibit externalizing or internalizing behaviors. Externalizing behaviors are observable behaviors, such as aggression, destruction of property, or impulsive behavior (e.g., yelling, shouting, etc.), that are closely associated with students with challenging behaviors. Internalizing behaviors, however, are more covert behaviors and, inherently, are more difficult to observe. Internalizing behaviors include symptoms of depression, sadness, or anxiety (Merrell & Walker, 2004). Both externalizing and internalizing behaviors can hinder the students' ability to make academic progress (Lane et al., 2008).

Mathematics and Students with Challenging Behaviors

A defining characteristic of students with challenging behaviors acknowledges that the student's behavior adversely impacts their academic performance (Federal Register, 2006). While students with challenging behaviors demonstrate underachievement in all academic subjects, students with severe challenging behaviors (i.e. EBD) experience the most difficulty with spelling and mathematics (Reid, Gonzalez, Nordness, Trout, & Epstein, 2004). For example, Epstein and colleagues (2005) noted that 42% to 93% of students with EBD demonstrated evidence of an MD, ranging between 1 to 2 years below expected grade level performance. Further, MD tended to persist over time for students with challenging behaviors.

Similarly, Nelson, Benner, Lane, and Smith (2004) conducted a cross-sectional study with a random sample of 155 students with challenging behaviors in kindergarten through 12th grade. From that study, the authors determined the overall mathematics performance of students with challenging behaviors decreased by 20 percentile points from childhood to adolescence. Wagner and colleagues (2004) described similar results, determining that mathematics calculation performance declined from the 34th percentile in elementary school to the 28th percentile in high school for students with challenging behaviors. This suggests that MD may worsen as students with challenging behaviors progress through their educational careers.

Wei, Lenz, and Blackorby (2012) also examined the mathematics growth and trajectories of all students with disabilities. Students in all disability categories, including students with EBD (i.e., students who exhibit challenging behaviors), displayed lower mathematics achievement levels than their general education peers (i.e., students without disabilities). In particular, students with EBD were shown to have similar mathematics growth trajectories as students with learning disabilities. Overall, students with challenging behaviors experienced slower mathematics growth when compared to their general education peers.

Mathematics Interventions

Interventions for students with MD. Previous reviews of the literature have examined mathematics interventions for students with MD. For example, Baker, Gersten, and Lee (2002) synthesized the mathematics literature for students with MD. The authors included 15 studies from 1971 through 1999 and ranged between 2nd and 11th grade. Baker et al. (2002) determined that four intervention approaches are beneficial for students with MD: (1) giving students and parents feedback on mathematics performance, (2) parent support, (3) peer-assisted learning interventions, and (4) utilizing explicit instruction.

Kroesbergen and Van Luit (2003) also conducted a synthesis of mathematics interventions for students with MD, with a focus on students in the elementary grades. The authors included 58 studies from 1985 to 2000. These researchers concurred that explicit instruction was extremely effective for students with MD in addition to interventions focused on foundational mathematics skill acquisition. More recently, Gersten et al. (2009) reviewed the literature to investigate interventions and instructional techniques to help students with MD. Again, researchers concluded that instruction during intervention should be explicit and systematic. Gersten et al. (2009) also determined that intervention material should provide students with visual representations of mathematical ideas and include approximately 10 min in each session to building fluency in mathematics facts. Specifically, for students that struggle with word-problem solving, strong evidence exists that interventions should be based on common underlying structures (Gersten et al., 2009). The most common underlying structures are called schemas.

Word-problem interventions. It is important for students to develop strong word-problem skill because word problems are how most students are asked to demonstrate mathematics competence. There are several approaches for teaching students with MD to solve word-problems. Those methods may include: (a) relying on key words to solve word-problems (Cutler & Monroe, 2006); (b) diagramming or drawing word problems (van Garderen, 2007); (c) metacognitive strategies (Case, Harris, & Graham, 1992); or (d) schema instruction (Fuchs et al., 2008; Powell, 2011; Riley & Greeno, 1988), to name a few. Schema instruction is a proven strategy for teaching word problems to students with MD. With schema instruction, students learn to categorize word problems based on schema, and students are taught strategies to solve word problems. Research on schema-based instruction has been found to be more effective than other strategies when teaching students with MD to solve word problems (Zhang & Zin, 2012).

Schema instruction. In schema-based instruction, students are taught to identify word-problems by the problem type and apply a specific strategy, tailored for that problem type, to solve the word problem. Surprisingly, almost all routine word problems that students see and solve in elementary and middle school fall into one of seven different schemas: total, difference, change, equal groups, comparison, combinations, or ratios and proportions. Contrasted with conventional word-problem instruction, students in schema instruction are taught to recognize a given word problem as belonging to a schema and employing strategies to solve that word problem. Research on schema instruction has highlighted the benefit of teaching this strategy to students with MD (Fuchs, Fuchs, & Prentice, 2004; Fuchs, Seethaler, et al., 2008; Jitendra, Griffin, Haria, et al., 2007).

Metacognitive strategies. Metacognitive strategies, incorporated with schema instruction, are also effective when teaching students with MD to solve word problems. As noted by Montague, Enders, and Dietz (2011), effective problem solvers use a variety of cognitive and metacognitive strategies to solve word problems such as identifying important information and disregarding extraneous information. Case et al. (1992) investigated the effects of using self-regulated strategy development on word problem solving for students with MD. Students were taught metacognitive skills and asked to apply those skills to addition and subtraction word problems. Several instructional procedures were employed including: preskill development, conferencing, discussion of problem-solving strategies, modeling of the strategy and self-instruction, mastery of the strategy, collaborative practice of the strategy, independent performance, and generalization of the strategy. Within the intervention, students were taught to identify key words, engage in discussions of progress, use a small chart to employ a problem-solving strategy, think alouds, and scaffolded instruction. The results indicated that all students showed improved performance on word-problem accuracy from baseline to intervention, especially in subtraction. Metacognitive strategies are most likely important to encouraging students to read the word problem and organize the word problem information in a meaningful way.

Mathematics Intervention for Students with Challenging Behaviors

Much of the research conducted on students with challenging behaviors (e.g., EBD) has concentrated on modifying behavior. Recently, however, several reviews of the literature have been conducted in an effort to uncover effective academic interventions for students with challenging behaviors (i.e., Hodge, Riccomini, Buford, & Herbst, 2006; Lane, 2004; Mooney, Epstein, Reid, & Nelson, 2003; Mulcahy, Krezmien, & Travers, 2016; Ralston, Benner, Tsai, Riccomini, & Nelson, 2014; Ryan, Pierce, & Mooney, 2008; Templeton, Neel, & Blood, 2008). In general, these reviews determined that several academic interventions also included components of self-determination. For example, much of the previous literature studied components of self-regulation, self-monitoring, or self-evaluation and used academic skills as a primary or secondary outcome. Overall, academic interventions in the literature were sparse with very few interventions targeting academic skill acquisition (i.e. reading, mathematics, and spelling).

Specific to mathematics, several syntheses for students with challenging behaviors have been conducted in recent years. For example, Hodge et al. (2006) reviewed instructional interventions in mathematics from 1985 through 2005. The authors located 13 studies that met their inclusion criteria, with participant ages ranging from 9 to 19 years old. Hodge et al. (2006) noted 77% of the studies located included a component of self-determination as the primary outcome (i.e., self-monitoring, self-regulated strategies, self-management, etc.) and mathematics as the secondary outcome. Specifically, 12 out of the 13 studies used mathematics calculation or basic facts as a dependent measure. Problem-solving interventions were almost non-existent for students with challenging behaviors. Hodge et al. (2006) concluded that further research on mathematics intervention for students with challenging behaviors was imperative.

With more of a focus on student characteristics, Templeton et al. (2008) conducted a meta-analysis of mathematics interventions for students with challenging behaviors examining four research questions. In general, Templeton et al. (2008) investigated how mathematics performance varies as a function of the student and participant characteristics. Templeton et al. (2008) also examined the levels of mathematics and the treatment, maintenance, and generalization effects of those interventions. Inclusion criteria included any mathematics intervention for students with challenging behaviors published in a peer-reviewed journal from 1976 through 2006. After applying inclusion criteria, 15 single-case studies were included in the meta-analysis, ranging from elementary aged students to secondary-aged students. Contrasted to the review conducted by Hodge and colleagues (2006), Templeton et al. (2008) included only two additional studies for the meta-analysis.

Templeton et al. (2008) concluded that mathematics performance did not vary as a function of study or participant characteristics. Of the mathematics instructional elements (e.g., strategy instruction, environmental accommodation, or instructional delivery), environmental accommodations (e.g., offering choices or feedback) was the only factor that provided a significant difference. Specifically, interventions without an element of environmental accommodations were surprisingly more effective. Maintenance probe results were disappointing while generalization probes appeared effective, although not often collected. Lastly, this meta-analysis discovered that more interventions used complex mathematics skills as a dependent variable rather than computation facts (e.g., word problems versus multiplication fact acquisition). Thereby, establishing a subtle shift from previous findings.

Ralston et al. (2014) identified and analyzed study components that were thought to be *best evidence* through the use of improvement rate differences and percentage of nonoverlapping data (PND). After the inclusion criteria were applied, 27 studies were utilized, ranging from 1968 through 2009. Ralston et al. (2014) concluded that there were three broad areas of mathematics research for students with challenging behaviors: (a) peer mediated strategies, (b) self-mediated strategies, and (c) teacher-mediated strategies. Teacher-directed mathematics interventions and instructional components contributed to treatment effects. For example, explicit teacher-directed instruction of difficult tasks was effective in improving the mathematics performance of two students with challenging behaviors (Lee, Sugai, & Horner, 1999).

Pursuant with preceding syntheses, Ralston et al. (2014) determined that the majority of the reviewed studies were self-regulation interventions, including self-evaluation, self-monitoring, or a combination of self-regulation interventions. Again, this synthesis concurred that the majority of the mathematical dependent measures were basic mathematics facts and operations, with very little to no research conducted on number sense, problem solving, or algebraic thinking. Ralston and colleagues (2014) again called for more higher-level mathematics research to be conducted and mathematics research for students with challenging behaviors that included more participants per study.

Most recently, Mulcahy et al. (2016) conducted a review of the literature examining mathematics interventions for students with challenging behaviors in secondary education settings. The synthesis included 19 articles, ranging from 1975 through 2012. The purpose of the review was to examine the methodological quality of the research on mathematics interventions for secondary students with challenging behaviors. The researchers applied and reported quality indicators for all studies using a combination of standards from Cook et al. (2015), Horner et al. (2005) and Kratochwill et al. (2010). The researchers reported significant variability in the quality of studies.

Furthermore, findings from Mulcahy et al. (2016) revealed that only eight studies actually investigated a functional relationship between mathematics interventions and mathematics performance for secondary students with challenging behaviors. The remainder of the studies investigated a functional relationship between behavioral or self-monitoring interventions and mathematics performance. Mulcahy et al. (2016) concluded that this was likely due to the common held belief that by intervening on challenging behaviors, academic performance would likely improve for students with challenging behaviors. However, as Mulcahy et al. (2016) discusses, this does little to help establish effective mathematical interventions for students with challenging behaviors. The researchers again reiterated that the majority of the interventions focused on foundational mathematical skills rather than advanced concepts especially needed for students in secondary education settings.

Word-Problem Interventions for Students with Challenging Behaviors

While all of the following interventions were included in the preceding syntheses, it is important to discuss specific word-problem interventions that have been previously researched on this population of students (i.e. students with challenging behaviors and MD). Alter, Brown, and Powell (2011) examined word-problem solving using the classroom teacher as the primary interventionist. All three participants were between first and fourth grade and had a primary school diagnosis of EBD. Two dependent measures were collected: the percentage of mathematics word problems solved and on-task behavior. The researcher created materials were designed for students to choose one of three strategies (i.e., Guess and Check, Draw a Picture, or Making a Table or Chart) to apply to the word problem. Teachers were trained during two workshops and instructed to reframe their methods of problem-based teaching to emphasize schemas. During implementation, the teacher utilized a script to implement the intervention. Results indicate that all students increased their percentage of word problems solved, however none of the participants were able to achieve over 50% of mathematics problems solved correctly.

Alter (2012) examined an eight-step checklist for solving word-problems for four students with challenging behaviors in grades 4 and 5. The problem-solving steps included: (a) Read the problem aloud, (b) Paraphrase, (c) Visualize, (d) State the problem, (e) Hypothesize, (f) Estimate, (g) Calculate, and (h) Self-check. Three dependent measures were collected: (a) percentage of problems completed correctly on daily worksheets, (b) percentage of problems completed correctly on a pre- and posttest measure, and (c) percentage of time on task. Results indicate that all students increased in their problem-solving accuracy, while 75% of the students increase in their rates of on-task behavior. Results were based on visual analysis of the multiple-baseline data.

Lastly, Peltier and Vannest (2016) investigated the effects of schema instruction on the word-problem solving on two fourth-grade students with challenging behaviors. A single-subject reversal design was used, and the dependent variable was the percentage of word-problems answered correctly. The intervention package consisted of schema instruction, strategy instruction, and the use of reinforcement for task completion. Students were taught three schemas: (a) total, (b) change, and (c) difference. The authors reported all students improved in accuracy from baseline. However, only two data points were taken during baseline and therefore, results should be interpreted with caution.

Word-Problem Intervention for Present Study

The intervention program used in this study is a schema intervention called Pirate Math. The original Pirate Math program was developed through a study funded by the National Institute of Child Health and Human Development and is considered an evidence-based intervention (Fuchs et al., 2010; Fuchs et al., 2009; Fuchs et al., 2008). Pirate Math was designed as a supplement to regular mathematics curriculum in addition and subtraction word problems. Broadly, Pirate Math helps teach elementary students with MD to focus on the conceptual framework of three types of word problems (i.e., schemas), identify extraneous information, set-up algebraic equations to mirror the structure of the given word-problem, solve for a missing variable, and self-evaluate the solution. Pirate Math also includes a standardized, token-based reward system. The version of Pirate Math used in this study is improved upon the original to include explicit instruction about setting up and solving equations that represent the word-problem schema.

The Pirate Math program incorporates several evidence-based components for students who struggle with mathematics. Specifically, (a) instruction is explicit and systematic, (b) word-problem instruction is based on common underlying schemas, (c) time dedicated in the intervention to fluency building, (d) progress monitoring is utilized, and (e) motivational strategies are included throughout the intervention (Gersten et al., 2009). In addition, there are several metacognitive strategies built into the scaffolded lessons of Pirate Math that help students remember to read the problem and organize the word-problem work.

Students begin each lesson by answering mathematics fact flash cards. Research suggests that fact fluency is an important skill for future success in mathematics and directly connects to the student's ability to perform more advanced mathematics (National Mathematics Advisory Panel, 2008). The Pirate Math intervention relies on schema instruction for solving word-problems, which is the third activity of every lesson. Schema instruction encourages students to recognize word problems as belonging to a problem type (i.e., schema) and to apply a strategy to solve that problem type. Schema instruction has been determined to be the most effective approach to word-problem solving (Xin & Zhang, 2009) and has been investigated in many research studies (e.g., Fuchs, Seethaler, et al., 2008; Griffin & Jitendra, 2009; Willis & Fuson, 1988). In addition to schema instruction, students are encouraged to use metacognitive strategies throughout Pirate Math, such as reading the problem, restating the problem, and thinking aloud to determine the schema. Lastly, students are also taught corresponding gestures for each problem type (i.e. total, difference, and change). Gesturing involves teaching the students hand-motions along with each problem-type equation. Previous research on gesturing indicates that students are able to deepen their understanding of mathematical concepts through gesturing (Novack, Congdon, Hemani-Lopez, & Goldin-Meadow, 2014) and therefore, is an integral part of the Pirate Math intervention.

In this chapter, I reviewed the previously published literature about students with mathematics difficulties and students with or at-risk for EBD. Students with MD historically struggle in mathematics but especially in word-problem solving. We established through a review of previously published research that students with challenging behaviors (e.g., EBD or at-risk for EBD) also have difficulty in mathematics and display patterns of mathematics achievement similar to students with MD. I also investigated the to determine an effective word-problem intervention. Pirate Math is multi-component, word-problem intervention. Little is known, however, about the effectiveness of the Pirate Math intervention for students that display challenging behaviors. In this next chapter, I discuss the methods utilized for the present study.

Chapter III: Method

This randomized control trial investigated three main research questions. First, I examined the differences in behavioral profiles (i.e., externalizing and internalizing behaviors) between students with MD and those without MD to determine whether more students with MD have behavioral profiles of concern. Second, I investigated the efficacy of a word-problem intervention for students with MD and whether behavioral profiles moderated student-level response to the word-problem intervention. Specifically, I asked the following research questions:

(1) Do students with MD have significantly different behavioral profiles when compared to students without MD? That is, do students with MD demonstrate a greater number of externalizing or internalizing behaviors compared to students without MD?

(2) What is the efficacy of a word-problem intervention for students with MD?

(3) Is the efficacy of the word-problem intervention for students with MD different based on their behavioral profiles? That is, do higher externalizing or internalizing behaviors predict student response to the word-problem intervention?

With regard to the first research question, I hypothesized that students with MD would have different behavioral profiles than students without MD, specifically demonstrating higher rates of at-risk behaviors versus students without MD. Previous research conducted by Epstein et al. (2005) noted that between 42% and 93% of students with EBD demonstrated evidence of a MD. Therefore, I would expect to find a similar pattern of co-morbidity within my sample population. For my second research question, I hypothesized that the word-problem intervention would be effective on students' word-problem solving. Previous iterations of this study have determined that Pirate Math is an effective intervention for students with MD versus their same-aged peers in the control group. Last, I hypothesized that behavior would predict response to the word-problem intervention (e.g., Pirate Math). Previous research has demonstrated that students that display at-risk behaviors may need more intensive and individualized intervention to support their academic and behavioral deficits (Maggin et al., 2016). Therefore, it is assumed that students that demonstrate high rates of at-risk behaviors would respond slower to the word-problem intervention when compared to their same-aged peers that do not present with at-risk behaviors.

Parent Study: Three-Year Efficacy Trial

The present study was conducted as part of a larger Goal 3 efficacy study funded by the Institute for Education Sciences (Powell & Barnes, 2015). The primary goal of the parent study was to test the efficacy of a tutoring intervention (e.g., Pirate Math) for students with or at-risk for MD and to understand the role of algebraic reasoning within problem solving. Data from three cohorts of third-grade students with MD were collected across three separate school years (i.e., 2015-2016, 2016-2017, and 2017-2018) to have enough statistical power to understand the efficacy of the intervention. This dissertation study (i.e., the present study) worked only with third-grade teachers and students from Year 3 of the parent study. Table 1 serves as a reference for distinguishing the parent study from the present study.

Table 1

Parent Study Versus Present Study

	Parent Study	Present Study
Participants	Third-grade students with MD and their classroom teachers	Third-grade students with MD and their classroom teachers
Goal	To test the efficacy of a word-problem intervention for students with MD	To examine behavioral profiles of students with and without MD and to learn whether behavior profiles moderate the efficacy of a word-problem intervention for students with MD
Study Duration	August 2015-April 2016 August 2016-April 2017 August 2017-April 2018	August 2017-February 2018
Measures	Addition (0-18) Double Digit Addition (0-18) Double Digit Subtraction (0-18) Equivalence Problems KeyMath 3 Numeration Mathematics Anxiety Rating Scale Open Equations Pennies Test Subtraction (0-18) Texas Word Problems WASI-II Matrix Reasoning WASI-II Similarities WASI-II Vocabulary Wisconsin Card Sorting WJ-IV Letter Pattern Matching WJ-IV Numbers Reversed WJ-IV Passage Comprehension WRAT Mathematics Computation WRAT 4 Word Reading WJ-IV Passage Comprehension WJ-IV Numbers Reversed	Open Equations Pennies Test Student Risk Screening Scale (SRSS) Texas Word Problems
Schemas Studied	Total Difference Change	Total Difference
Total Sessions	52 sessions	30 sessions

Participants

Teachers. In this study, we recruited third-grade general education teachers, and the students in the teacher classrooms were the participants in the study. After gaining permission from the Austin Independent School District's office of external research, we contacted elementary school principals to gauge their interest in the study. If principals agreed, we recruited third-grade general education classroom teachers in the principals' schools. Teachers had to consent to participation in the study by signing and dating a teacher consent form. In the present study, teachers were asked to fill out two measures on all the students in their classroom: (a) student demographic information and (b) the Student Risk Screening Scale (SRSS; Drummond, 1994). By the conclusion of the study, there were 39 general education teachers consented for participation. Classroom teachers received $50 gift cards for their participation in this study.

Students. To be eligible for the study, students must have been a current third-grade student enrolled in a public elementary school within the Austin Independent School District and in the classroom of a participating teacher. Before the study began, parents or guardians consented for their child's participation in the study, and students provided assent for study participation.

Our focus for the present study was third-grade students with MD. MD was defined as performance below the 13th percentile on a test of single-digit word-problems (e.g. Pennies Test; Jordan & Hanich, 2000). The 13th percentile was determined based on recruitment and assessment of third-grade students in the first year of the parent study. In addition to scoring below the 13th percentile, students also had to demonstrate minimal

English understanding and conversational skill. Participating students need to demonstrate proficiency in English due to the tutoring being conducted solely in English. Table 2 provides the demographics of the study participants.

Table 2

Overall Participant Demographics

	Overall		Typical		MD	
	n	%	*n*	%	*n*	%
Gender						
Male	309	50.20%	247	53.20%	62	41.05%
Female	306	49.80%	217	46.80%	89	58.94%
Race						
African American	78	12.67%	57	12.30%	21	13.91%
Asian	22	3.57%	18	3.90%	4	2.65%
Caucasian	139	22.56%	135	29.10%	4	2.65%
Hispanic/Latino	331	53.80%	220	47.40%	111	73.51%
Other	45	7.30%	34	7.30%	11	7.29%
ELL Status	265	43.10%	170	36.60%	95	62.91%
SPED Services						
Yes	50	8.10%	30	6.50%	20	13.25%
No	565	91.90%	434	93.50%	131	86.75%

Note. ELL = English Language Learner; SPED = special education

Tutors

Graduate research assistants (GRA) conducted all screening, pretesting, intervention, and posttesting. All sessions were audio recorded for later analysis of procedural fidelity. Between 13 and 15 GRAs were utilized to conduct all sessions. Each GRA has or was seeking a Master's or doctoral degree in an education-related field. All GRAs received over 20 hours of training in order to conduct screening, pretesting, intervention, and posttesting. GRAs also completed reliability checks with the project manager for any measure that was administered to students.

Research Design

The parent study was a blocked, randomized control trial that examined the impact of a word-problem intervention (i.e., Pirate Math) on word-problem solving outcomes for students with MD. Randomized control trials (RCTs) are high-quality studies in which participants are assigned to two or more groups to determine the effects of an intervention (Institute of Education Sciences, 2014). In the parent study, qualifying students were randomly assigned, blocking by school and classroom, to one of three conditions: (a) control (i.e., no tutoring), (b) word-problem tutoring without equal sign instruction, and (c) word-problem tutoring with equal sign instruction. In the present study, we utilized this research design and analyzed the word-problem performance and behavioral profiles of students in the screening pool (i.e., all third-grade students) and students randomly assigned to one of the three conditions.

Materials and Procedures

Tutoring. The word-problem intervention that is the focus of the present study is a program called Pirate Math. The Pirate Math intervention of the parent study included over 50 one-on-one lessons, implemented three times a week, with each session lasting about 30 min. For the present study, all students completed at least Lesson 30 of the

Pirate Math Intervention before they were posttested. The parents study incorporates two treatment groups in which the effects of explicitly teaching the equal sign plus Pirate Math are compared to students only receiving the Pirate Math intervention. For the purposes of the present study, the data was not disaggregated amongst the two tutoring conditions. The tutoring occurred during the school day at a time agreed upon with the teacher and the GRA. Tutoring took place over approximately 15 weeks. GRAs administered each lesson. Each lesson consisted of five components led by the GRA: (a) mathematics fact flashcards, (b) algebraic-reasoning activity or mathematics concepts activity, (c) explicit instruction on solving word problems, (d) word-problem sorting practice, and (e) cumulative review. Table 3 provides a general outline of the Pirate Math lessons in the present study. Table 4 provides an overview schedule of the Pirate Math intervention in the present study, from screening to posttesting.

Table 3

Pirate Math Unit and Lesson Overview

Unit	Days	Topics Covered
Intro	1-4	• Solve basic addition problems, with and without regrouping • Solve basic subtraction problems, with and without regrouping • Label graphs (i.e., bar graphs, pictographs, pictographs with multiplier, tables) • Introduce and discuss meaning of the equal sign
Total	5-16	• Introduce Total problems (Part 1 + Part 2 = Total) • Solve Total problems with Total missing (e.g., 5 + 4 = X) • Solve Total problems with one Part missing (e.g., 5 + X = 9) • Introduce three-part Total problems (Part 1 + Part 2 + Part 3 = Total) • Solve three-part Total problems with Total or one Part missing (e.g., 5 + 4 + 2 = X; 5 + X + 2 = 11) • Solve Total problems with graphs (i.e., bar, pictographs, pictographs with multipliers, tables) • Use cubes to balance both sides of the equal sign • Use cubes to solve missing addend problems • Draw pictures to solve missing addend problems • Draw pictures to balance both sides of the equal sign • Solve equations by balancing sides and isolating the X • Solve Total equations by balancing sides • Practice computation (addition and subtraction with and without regrouping)
Difference	17-33	• Introduce Difference problems (Greater – Less = Difference) • Introduce and practice identifying compare words (e.g., more, fewer, less, taller, smaller, faster) • Solve Difference problems with Difference missing (e.g., 8 – 4 = X) • Solve Difference problems with number that's greater missing (e.g., X – 4 = 4) • Solve Difference problems with number that's less missing (e.g., 8 – X = 4) • Solve Total problems (i.e., Total missing, part missing, three-part Total problems with Total or part missing) • Solve Total and Difference problems with graphs (i.e., bar graphs, pictographs, pictographs with multipliers, tables) • Use cubes to balance both sides of the equal sign with subtraction problems • Solve Total and Difference equations by balancing sides • Draw pictures to solve missing addend, minuend, and subtrahend problems

Table 4

Timeline for Pirate Math Implementation

Month	Activity
August-September	• Teacher and parent consents and student assents • GRA training • Screening of third-grade classrooms
September-October	• Pretesting of third-grade students with MD • Behavior measure (i.e., SRSS) completed by teachers • Fidelity checks (on-going through January)
October-January	• Intervention implementation
January-February	• Posttesting

Description of Intervention Activities

In the next five sections, I describe the daily activities that occur as part of the Pirate Math intervention. GRAs conducted each of these activities with an individual student, and the GRA for each student remained the same during intervention implementation. In this section, we refer to the GRAs as tutors.

Activity #1: Mathematics fact flashcards. GRAs started every Pirate Math session with a fluency-building activity in which students were presented with addition and subtraction flashcards. Addition flashcards featured addends of 0 to 9. The subtraction flashcards featured minuends of 0 to 18 and subtrahends from 0 to 9. On Days 1 through 4 of Pirate Math, the tutor taught the student counting-up strategies for learning addition and subtraction facts. The student was prompted to implement each of the counting-up strategies throughout the lessons (Tournaki, 2003). In the flashcard activity, the student had 1 min to complete each of the two trials. After the initial 1 min trial, the tutor and student counted the number of flashcards answered correctly. The tutor also provided immediate, corrective feedback to the student by reviewing the counting-up strategies for any noted errors. Prior to starting the second 1 min trial, the tutor challenged the student to beat their previous score. This encouragement was included as a way to alleviate the tedious nature of the fluency building activity and motivate the student to improve their score (Fuchs et al., 2008). At the end of the second 1 min timing, the tutor and student graphed the highest score (see Figure 1 for a sample graph). The graph served as a self-regulation tool for setting future goals and monitoring student progress (Wehmeyer, Sands, Doll, & Palmer, 1997).

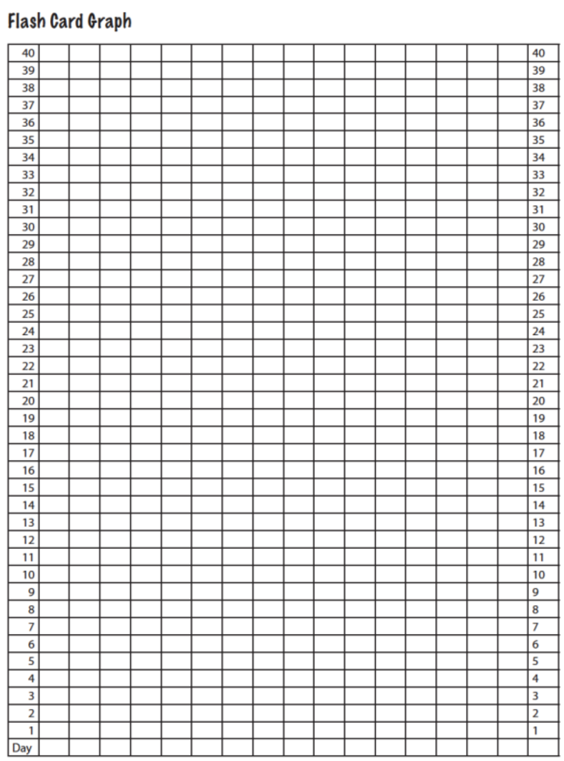


Figure 1. Flashcard graph.

Activity #2. In the parent study, students in the word-problem intervention groups were randomly assigned to one of two options for the second activity of each tutoring lesson. A description of the two activities are provided in order to be transparent about all of the intervention materials. In the present study, however, the two groups were collapsed at final analysis to only evaluate the treatment groups versus the control group.

Activity #2a: Equation quest. Students in the word-problem intervention with equal sign instruction group learned the relational meaning of the equal sign through an activity called Equation Quest. During this 2 min activity, the tutor reintroduced the equal sign (=) and taught the student to understand the meaning of the equal sign as *the same as*. Through a sequence of activities, starting at Lesson 3, the student learned that the equal sign acts as a balance between two sides of an equation and does not solely signal a calculation.

In the initial lessons, the tutor and student worked with manipulatives (i.e., cubes) to determine if the number of cubes on the left side of the equation was *the same as* the number of cubes on the right side of the equation (see Figure 2 for sample problems). The tutor prompted the student to add or subtract cubes from one side of the equation to make both sides *the same*. As the lessons progressed, the tutor and student drew pictures or shapes to represent and solve equations. Ultimately, the work with manipulatives and drawings was instigated to help students solve equations in the abstract form (i.e., with numbers and symbols). Students were tasked with determining whether the left side of the equation was *the same as* the right side of the equation. The students in this intervention condition learned a sequence of steps to balance and solve an equation. To provide consistency across lessons and student participants, the tutor maintained the language of the equal sign (e.g., *the same as*) throughout the Pirate Math lessons for all students in the word-problem and equal sign intervention.

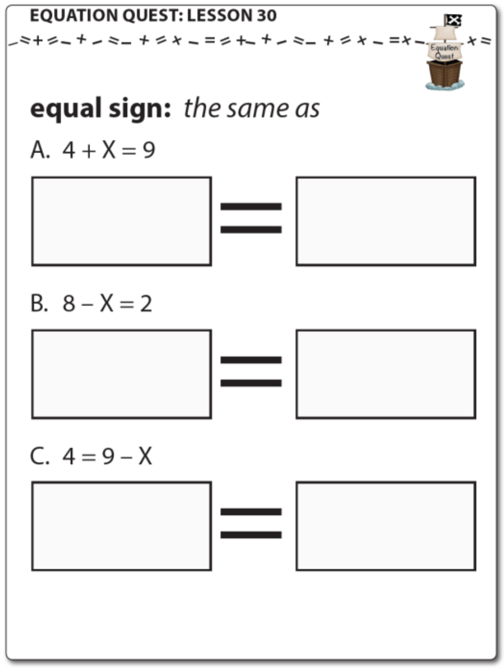


Figure 2. Sample Equation Quest activity.

Activity #2b: Pirate crunch. Students that received the word-problem intervention without equal sign instruction completed a 2 min activity called Pirate Crunch. Pirate Crunch lessons began on Lesson 3. During Pirate Crunch, students reviewed previously learned mathematics material (see Figure 3) not related to word problems or the equal sign. Specific mathematical topics included: mathematics vocabulary, perimeter, area, fractions, money, order of operations, and telling time.

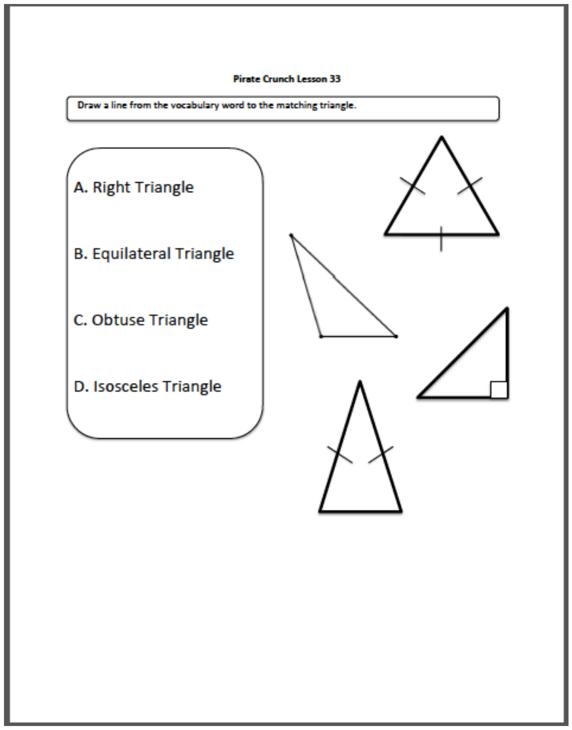


Figure 3. Sample Pirate Crunch activity.

Students in the word-problem intervention without equal sign instruction condition did not receive explicit instruction on the equal sign. Pirate Crunch was included to control for intervention time so the minutes of instruction in both active word-problem intervention conditions did not vary significantly.

Activity #3: Buccaneer problems. The third activity consisted of tutor-led explicit instruction through an activity called Buccaneer Problems. Days 1 through 4 of the Buccaneer Problem lessons included a review of addition and subtraction skills using the *counting up* strategies to assist with mathematics fact difficulties and regrouping for double-digit calculations. Students also learned how to label and interpret data presented on charts and graphs. On Days 5 through 16, the student began to solve Total word problems. Tutors introduced Difference problems on Day 17. Each Buccaneer Problem activity included a review of previously learned material, was scaffolded through tutor-led questioning, and required the student to verbalize their thinking.

On Day 5, students were taught a specific attack strategy that relies on metacognitive strategies and provided a visual prompt (see Figure 4) called *RUN*. Students learned to think through the problem before solving it by first checking for a table or a graph and numbering it (if applicable), and then *RUN*ning through the problem. To *RUN*, the student had to *R*ead the problem, *U*nderline the label and cross out irrelevant information, and *N*ame the problem type (i.e., choose the correct schema to use) by asking questions about the problem. Students were taught prompts to determine the problems types. For example, to decide if a problem was a Total problem, students asked themselves: *Are two or more parts being put together for a total?* For Difference problems, the student asked: *Are two amounts being compared for a difference?* The student then employed a series of written steps to solve the problem. The steps are specific to each of the schemas, but for all problem types, the students were taught to use an equation to represent the problem and to mark “X” to represent the missing information. Students were also taught hand gestures for each problem type and were encouraged to use those hand gestures throughout the lessons.

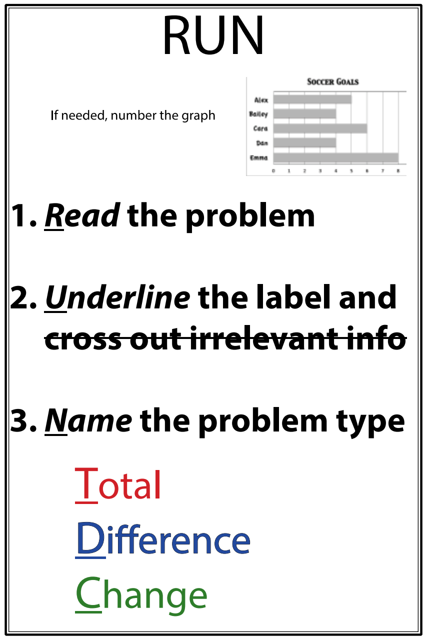


Figure 4. RUN Strategy Poster.

Total problems. As stated, the Total problem schema was first introduced on Day 5. In Total problems, students have parts that are put together for a total. The missing information from the word problem (i.e., "X") could be the total or one of the parts. After students identified word problems as fitting into the Total schema, students used an equation that represented the Total schema to organize word-problem information (i.e., P1 + P2 = T; where P is a part and T is the total). After checking for a table or a graph and using the attack strategy *RUN*, the student used five steps to find the solution of Total problems:

(1) Write P1 + P2 = T

(2) Find T

(3) Find P1 and P2

(4) Write the signs

(5) Find X

When these steps were combined with the equation-solving skills learned in Equation Quest, the student's paper looked similar to Figure 5. Students in the Word Problem only group solved for "X" by adding or subtracting in any way that made sense for the students. For Total problems with more than two parts, the student was taught to simply alter the equation (i.e., P1 + P2 + P3 = T) and solve according to the five Total steps.

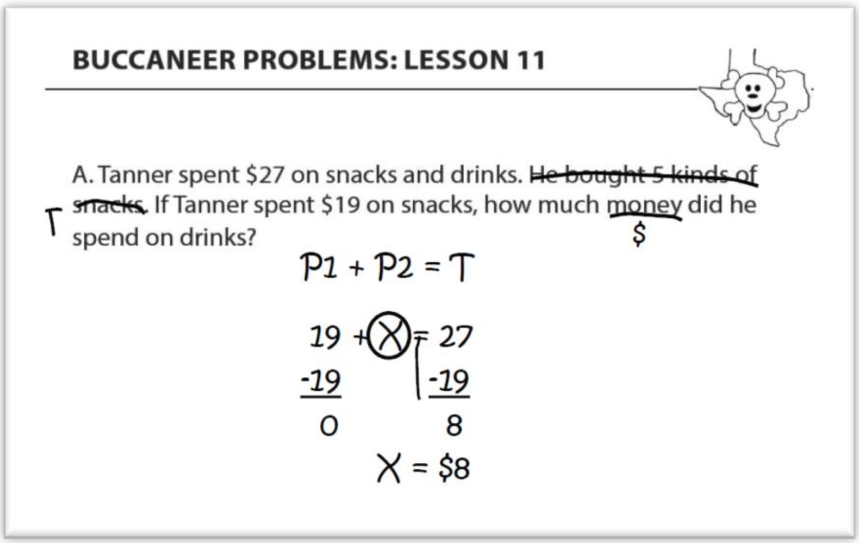


Figure 5. Total problem example.

Difference problems. Difference problems were introduced on Day 17. In Difference problems, the student learned to compare an amount that is greater and an amount that is less to find the difference. The missing information (i.e., "X") for Difference problems could have been the amount that is greater, the amount that is less, or the difference. To successfully identify the greater amount, the lesser amount, and the difference, the tutor taught the student to locate a "compare sentence." This compare sentence featured a comparison (e.g., *more*, *less*, or *fewer*). There were six steps to solving a difference problem:

(1) Write G – L = D

(2) Put brackets around the compare sentence and label G and L

(3) Find D

(4) Find G and L

(5) Write the signs

(6) Find X

As with the Total problems, students that were in the word-problem intervention plus equal sign instruction were taught how to solve for "X" by isolating the variable and balancing around the equal sign (see Figure 6 for an example). Students in the word-problem intervention without equal sign instruction determined whether to add or subtract to solve the problem without a focus on the equal sign.

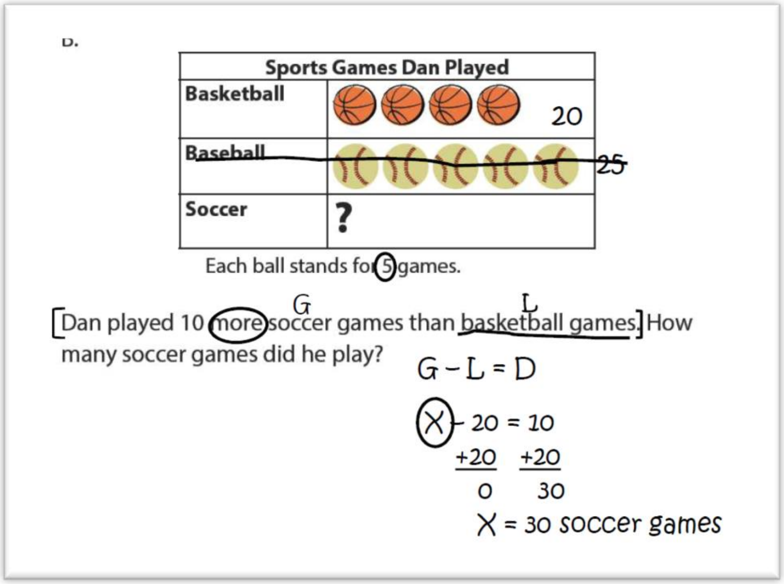


Figure 6. Difference problem example.

Activity #4: Shipshape sorting. Shipshape Sorting was a 1 min timed activity that allowed the students to practice identifying word-problem schemas (i.e., Total or Difference) learned during the Buccaneer problems. Before the sorting activity began, a mat with four squares was placed in front of the student. Each square was labeled with a word-problem type letter (i.e., T for Total, D for Difference, or C for Change) or the question mark symbol (see Figure 7). (Please note that, in the present study, students only identified Total or Difference problems and ignored the Change problem type.) The tutor reviewed the two word-problem schemas and explained the directions to place each word-problem card on the square with the corresponding problem type letter (i.e., T, D, or ?). The tutor reminded the student to sort the word-problem cards only and not solve any of the word problems. The tutor read all word-problem cards aloud. After 1 min, the tutor reviewed the sorting cards on the mat by providing immediate, corrective feedback to the student on at least three of the word-problem cards. Shipshape Sorting activity provided students the opportunity for increased practice identifying word-problem schemas.

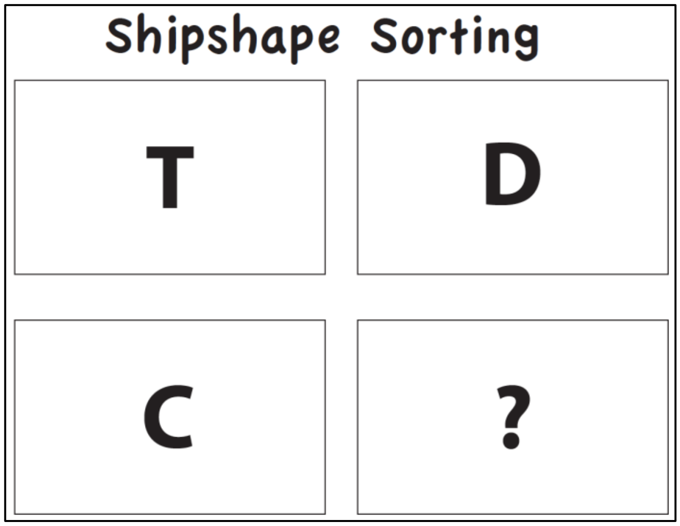


Figure 7. Sorting mat.

Activity #5: Jolly roger review. The Jolly Roger Review included a brief, timed review of the lesson contents. In the first activity, the student had 1 min to answer up to nine mathematics problems (i.e., single and double-digit addition or subtraction problems) or write appropriate equations for the two word-problem schemas (e.g., Total equation: P1 + P2 = T). Next, the student had 2 min to complete a word-problem using the appropriate schema steps taught during the Buccaneer Problem activity (see Figure 8). The student performed the timed review autonomously and then received feedback from the tutor, which reinforced content mastery.

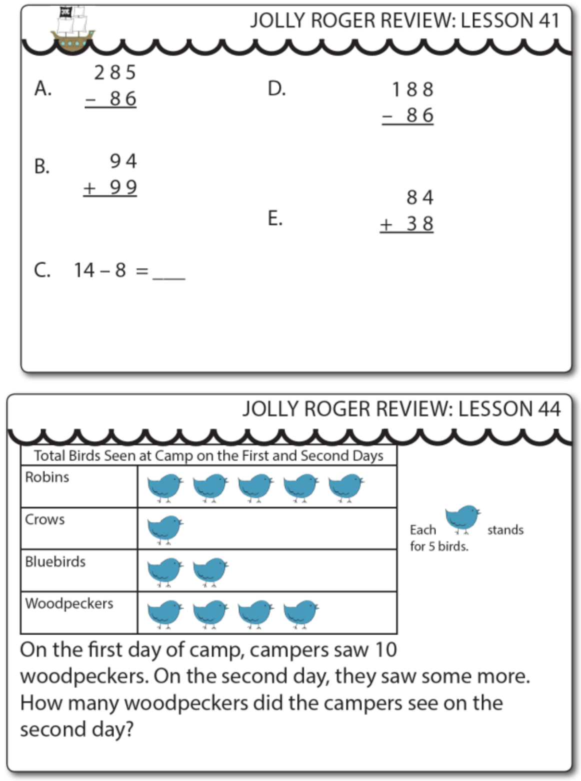


Figure 8. Sample Jolly Roger Review.

Motivation. The Pirate Math intervention incorporated a standardized, token-based reward system. At the beginning of each lesson, the tutor reviewed the Pirate Math rules (see Figure 9) with the student. When students followed the Pirate Math rules, they received gold coins throughout the lesson. Typically, students earned between three to six coins per lesson. At the conclusion of the lesson, the student recorded the number of coins they received that day by coloring in their treasure map. When the student reached the treasure box, they received a prize. Students typically earned one treasure chest prize per week.

Figure 9. Pirate math rules.

Control Group

Students in the control group were assigned to not receive the intervention. Students received mathematics instruction as would normally occur in a third-grade classroom. Information was collected from teachers via a survey about the word-problem instruction that students in the control group received in order to gain a better understanding of what a normally occurring third-grade mathematics class looked like for students in the control group. Out of the 39 teachers that were surveyed, 35 teachers (90%) utilize *Go Math!* as their core mathematics curriculum. *Go Math!* is a program created by Houghton Mifflin Harcourt publishing and is a comprehensive mathematics program available across kindergarten through 6th grade. Teachers were also queried about specific word-problem strategies. Overall, 29 teachers (74%) reported using a metacognitive attack strategy (e.g., CUBES, WIK, WINK, UPS Check or STAR) to teach word-problems. Only one teacher reported using schema instruction to teach students word-problem solving, and this teacher only mentioned one schema (i.e., Total). Students in the control group participated in the project by completing all pre- and posttesting sessions. Demographic and behavioral data was collected on students from the control group.

Data Collection and Measures

Student Risk Screening Scale (SRSS). The SRSS (Drummond, 1994) is a brief screening tool used to assess students' at-risk levels for externalizing and internalizing behaviors (see Appendix A). The SRSS included 12 items and required approximately 15 min for all teachers to rate all of the students in their classroom. It was recommended that the teachers have known the students approximately 6 to 8 weeks before filling out the SRSS, and we adhered to this recommendation as teachers provided this information in late September or early October, approximately 6 weeks into the school year. The SRSS measure is most effective without providing operationalized definition for each of the risk factors, and, therefore, we did not provide specific definitions for each of the 12 items. The items were scored on a 0 to 3-point Likert-type scale, where 0 indicated *never* and a score of 3 indicated *frequently*. Subsequently, total scores that are higher indicate higher at-risk status for externalizing or internalizing behaviors.

On the SRSS, the first seven items asked teachers to respond to questions relating to the frequency of the following externalizing behaviors: (a) steal; (b) lie, cheat, sneak; (c) behavior problem; (d) peer rejection; (e) low academic achievement; (f) negative attitude; and (g) aggressive behavior. For externalizing behaviors, the following scores were recommended to determine a student's at-risk status: low (0-3), moderate (4-8), or high (9-21; Drummond, 1994).

Questions 8 through 12 of the SRSS specifically asked about the frequency of the following internalizing behaviors: (a) emotionally flat; (b) shy, withdrawn; (c) sad, depressed, (d) anxious; and (e) lonely. For internalizing behaviors, the following were the recommended cut-scores at the elementary level: low (0-1), moderate (2-3), or high (4-15; Lane et al., 2015).

Many studies have been conducted examining the reliability and validity of the SRSS. Researchers have found high internal consistency (α = .81-.82) and test-retest reliability, where $r = .86$ (Oakes et al., 2010). Convergent validity exists between the SRSS and the Strengths and Difficulties Questionnaire (Lane, Parks, Kalberg, & Carter, 2007), the Systematic Screening of Behavior Disorders (Lane, Little, et al., 2009), and the Teacher Report Form (Lane, Oakes, et al., 2016). The SRSS has been validated with preschool students (Lane, Oakes, & Menzies, 2015); elementary students (Lane, Oakes, & Harris, 2012; Lane, Oakes, & Swogger, 2015); middle-school students (Lane, Oakes, & Carter, 2013); high school-aged (Lane, Robertson, & Parks, 2008; Lane, Oakes, & Ennis, 2013; Lane, Oakes, & Cantwell, 2016); students in rural elementary schools (Lane, Menzies, & Oakes, 2012); students in suburban elementary schools (Menzies & Lane, 2012); students in urban elementary schools (Oakes, Wilder, & Lane, 2010; Ennis, Lane, & Oakes, 2012); students in urban middle schools (Lane, Bruhn, & Eisner, 2010); students who are English learners (Lane, Richards-Tutor, & Oakes, 2014); and for use in functional behavior assessments (Clinton, 2010).

In addition to the SRSS, classroom teachers in the present study completed a survey about the extent of the behaviors students exhibit in their classrooms. Specifically, teachers were asked to identify students in the word-problem study (a) that received Individualized Education Plan (IEP) services under the classification of EBD, (b) students that received behavioral supports through a 504 plan, or (c) students that receive behavioral intervention through multi-tiered systems of support (MTSS). In this sample, teachers reported that three students (2%) had IEPs for EBD and were receiving specially designed instruction for their behavior, two students (1%) were receiving behavioral support through a 504 plan, and nine students (6%) were receiving behavioral support through a multi-tiered system of support framework. Students that were receiving behavioral support (i.e., IEP, 504 or MTSS) were equally distributed between the treatment and control groups.

Screening measure. The Pennies Test (Jordan & Hanich, 2000) was used as the screener to identify students with MD, and this test was administered before the other pretests. This test assessed the students' ability to solve 14 one-step addition or subtraction word problems. For each item, the GRA read the word-problem prompt aloud and provided time for students to solve the problem and write an answer. GRAs could read the problem aloud once more (for a total of two readings) upon student request. Based on Year 1 results from the parent study, Cronbach's alpha was .87.

Pretesting measures. The pretesting battery consisted of three assessments: Pennies Test, Texas Word Problems, and Open Equations. We described Pennies Test in the section about Screening because it was utilized as the determinant for MD status, as well as the pretest measure.

On the Texas Word Problems (Powell, 2015), we asked students to solve several types of additive word problems: Total, Difference, and Change. For the present study, we analyzed student performance on Total and Difference problems, with one Change problem to analyze possible carryover effects. Overall, there were three Total problems, three Difference problems, and one Change problem with single- or double-digit computation. GRAs read each word problem aloud and provided time for students to solve the problem and write an answer. Upon student request, each problem could be read an additional time. Based on Year 1 from the parent study, Cronbach's alpha was .88.

Open Equations (Powell, 2007) required the student to solve standard (e.g., 4 + __ = 5) and nonstandard (e.g., 4 + 3 = __ + 2) equations within 6 min. Cronbach's alpha for the Open Equations test was .94, based on data from Year 1 parent study.

At pretest, all three measures were administered to groups of third-grade students. Each GRA followed a pretesting protocol and was expected to read the protocol verbatim.

Posttesting measures. At the conclusion of at least Lesson 30 (i.e., after completion of Units 1-3), all students were administered a posttest. The posttest battery (see Appendix C) included one 30 min assessment that consisted of the Pennies Test, an abbreviated Texas Word Problem test with Total, Difference, and one Change problem from the full Texas Word Problems, and the Open Equations test. All assessments mirrored the pretesting measures. The Pennies Test and Texas Word Problems were untimed with the GRA reading each problem aloud, and Open Equations was timed for 6 min. At posttest, all three measures were administered to individual students or small groups of students. Each GRA followed a posttesting protocol and was expected to read the protocol verbatim.

Fidelity

Fidelity data was collected at screening, pretest, intervention, and posttest. Procedural fidelity checks were assessed two ways: (a) direct observation and (b) digital audio recording. The project manager conducted direct observation fidelity checks on 30% of tutoring sessions. Fidelity was also assessed through the use of a fidelity checklist (see Appendix D for an example fidelity checklist). Fidelity was considered acceptable at 90%. When fidelity was assessed below 90%, subsequent direct observations occurred until 90% was achieved. The overall fidelity of direct observations ranged between 79% to 100% with an average of 97.08%. Fidelity was also evaluated through audio recordings. By the conclusion of the parent study, GRAs will have listened to 20% of pretest, posttest, and tutoring sessions and determined the procedural fidelity using a checklist similar to Appendix D.

Chapter IV: Results

This study investigated the effects of behavior and a mathematics word-problem intervention. First, attrition analyses were conducted to determine that there were no systematic biases due to differential attrition. Preliminary analyses were completed for Research Question 2, and main analyses were conducted for all Research Questions. ANOVA models were used to determine the significance of effects, if any. Last, a regression model was used to answer Research Question 3.

Attrition Analysis

There were 819 students screened at the beginning of 2017-2018 for the parent study. There were 230 students that scored below the 13th percentile on the Pennies Test. Of these, 58 were initially disqualified due to the following reasons: behavioral issues (n = 7), did not complete full screening battery (n = 5), limited English status (n = 23), parent or student opted out of participation (n = 5), student moved schools (n = 4), too many students qualified in the same class (n = 7), severe disability status (n = 1), and students receiving special education who were already being pulled out frequently throughout the day (n = 6). During pretesting, an additional 11 students were lost for the following reasons: missing consents (n = 7), additional behavioral challenges (n = 2), limited English proficiency (n = 1), and withdraw from school (n = 1). A total of 162 students were initially randomized into one of two treatment groups or the control group, specifically, 98 students in the tutoring groups and 64 students in the control group. By the end of posttesting, 151 (93%) students remained. Of the 11 students who left the study, nine moved schools, one went into protective custody, and one was suspended for

30 days. From the 11 students, two students left the study from the control group and nine students from the treatment groups. What Works Clearinghouse (WWC) criteria were used to calculate overall and differential attrition (Institute of Education Sciences, 2014). Overall attrition was calculated at 7% while differential attrition was 6%. According to WWC, the combination of overall and differential attrition is expected to result in an acceptable level of bias even under conservative assumptions.

Main Analysis: Research Question 1

My first research question investigated the behavioral profiles of students with MD when compared to students without MD? That is, do students with MD demonstrate a greater number of externalizing or internalizing behaviors compared to students without MD?

Table 5

Descriptive Statistics for Externalizing and Internalizing Behavior

	MD	Not MD
Measure	Mean (SD)	Mean (SD)
Externalizing	3.73 (3.42)	2.01 (3.18)
Internalizing	1.77 (2.25)	0.84 (1.48)

A total of 441 student behavioral profiles were collected from the overall sample of 615 students. Four teachers failed to return completed SRSS rating scales, accounting for missing data of approximately 151 students. In addition, 14 students moved, seven students did not return consents, one student was long-term suspended, and one student entered into protective custody. Table 5 shows the descriptive statistics for externalizing and internalizing behaviors delineated by students with MD and students without MD. According to the SRSS, higher scores are consistent with higher at-risk status.

In order to determine if students with MD exhibited statistically significant higher rates of externalizing and internalizing behaviors compared to students without MD, a one-way ANOVA analysis was completed for each behavior. If the *p*-value was less than 0.05, then the difference was considered statistically significant. Table 6 reports the ANOVA results, including the *F*-value, *p*-value, and the Cohen's *d* effect size.

Table 6

ANOVA Results and Effect Sizes for Externalizing and Internalizing Behaviors Comparing Students With and Without MD

	F	*p*-value	Cohen's *d*
Externalizing	25.71	*<.001	0.52
Internalizing	26.08	*<.001	0.49

Note. * = statistically significant at p <.05

Results indicate that there was a significant difference in externalizing and internalizing behaviors between students with MD and without MD. That is, students with MD displayed higher externalizing and internalizing behaviors than students without MD. The Cohen's *d* effect size for externalizing behaviors, $F(1, 439) = 25.71$, $p < .001$, indicated a medium effect size, while the effect size of the internalizing behaviors, $F(1, 439) = 26.08$, $p < .001$, indicated a small to medium effect size.

Preliminary Analysis: Research Question 2

For the second research question, I investigated the effects of a word-problem intervention on students with MD. First, preliminary analyses were conducted to ensure assumptions for ANOVA were not violated (i.e., inspecting the data for abnormalities, ensure normality of distribution in the dependent variable, and to determine homogeneity of variances). Additional chi-square analyses were conducted on demographic data as well as reliability of the outcome measures.

All assumptions for ANOVA were not violated. There were no abnormalities in the data and the dependent variables were normally distributed. The test of Homogeneity of Variances was also conducted to ensure that all comparison groups have the same variance within the data. This assumption of ANOVA was not violated either.

The chi-square analysis was conducted on all demographic information to test for the "goodness of fit" between observed and expected between the treatment and control groups. The chi-square analysis yielded no significant differences in sex, race, English Language Learner status, or special education status between groups.

I also determined reliability of each of the measures described in the study. The Pennies Test consisted of 14 items (α = .724), the Texas Word Problems test consisted of seven-word problems (α = .828), and Open Equations consisted of 30 items (α = .877). All tests were considered within the "acceptable" limits for reliability.

Main Analysis: Research Question 2

My second research question investigated the effectiveness of the word-problem intervention for students with MD compared to the control group using three outcome measures: Pennies Test, Texas Word Problems, and Open Equations. Table 7 and Table 8 contains the descriptive statistics for the pre-test and unadjusted posttest measures. The means were not statistically significant at pre-test. Additionally, adjusted post-test means were generated but were not vastly different from unadjusted post-test means. Therefore, unadjusted posttest means were used in subsequent analyses. In order to assess the difference between pre and post-test, gain scores were calculated for each outcome measure (Table 9) and ANOVAs analyses were used to determine significance (Table 10).

Table 7

Descriptive Statistics for Pretest Measures

	Comparison	Treatment
Measure	*M* (*SD*)	*M* (*SD*)
Pennies Test	4.98 (1.74)	4.75 (1.74)
Texas Word Problems	2.10 (1.60)	2.19 (1.60)
Open Equations	5.69 (3.81)	5.01 (4.00)

Table 8

Unadjusted Means at Posttest

	Comparison		Treatment	
Measure	*N*	*M* (*SD*)	*N*	*M* (*SD*)
Pennies Test	62	7.68 (3.19)	89	7.67 (3.41)
Texas Word Problems	62	2.58 (2.47)	89	5.37 (3.72)
Open Equations	62	9.76 (6.06)	89	10.25 (6.00)

Table 9

Gain Scores at Posttest

	Comparison		Treatment	
Measure	*N*	*M* (*SD*)	*N*	*M* (*SD*)
Pennies Test	62	2.69 (3.34)	89	2.92 (3.36)
Texas Word Problems	62	0.48 (2.42)	89	3.18 (3.43)
Open Equations	62	4.06 (5.15)	89	5.24 (5.83)

Table 10

ANOVA Results and Effect Sizes

	F	*p*-value	Cohen's *d*
Pennies Test	0.12	.68	0.07
Texas Word Problems	28.38	*<.001	1.11
Open Equations	1.62	.21	0.21

Note. * = statistically significant at $p < .05$

Neither the Pennies Test, $F(1,149) = .17$, $p = .07$, nor Open Equations, $F(1, 149) = 1.62$, $p = .21$, showed statistical significance. The Texas Word Problems, $F(1, 149) = 28.38$, $p < .001$), however, was statistically significant, demonstrating a large effect size. In other words, students in the treatment groups significantly outperformed the control group in Texas Word Problems gains from pre- to posttest.

Main Analysis: Research Question 3

My third research question investigated the efficacy of the word-problem intervention based on student behavioral profiles. That is, do higher externalizing or internalizing behaviors predict student response to the word-problem intervention? Before calculating any statistics, a code was developed to indicate when students were considered high at-risk for externalizing or internalizing behaviors. The SRSS considers a student a high at-risk for externalizing behaviors at elementary if a student scored nine or above on the externalizing items (i.e., steal, lie, cheat, etc.) and high at-risk for internalizing behaviors if a student scored four or above.

Externalizing. Table 11 shows the descriptive statistics for student gains across all outcomes for students in the treatment group for students observed high at-risk versus not high at-risk in externalizing behaviors. Most notably, Texas Word Problems and Open Equations observed the greatest different in mean scores between students deemed high at-risk versus students not high at-risk of externalizing behaviors. However, in subsequent analysis using ANOVA (Table 12), only Texas Word Problem gains were statistically significant with, $F(1, 75) = 4.24$, $p = 0.04$, and a Cohen's d of 1.01 indicating large effect size. Table 13 displays the descriptive statistics for students in the control group with high at-risk versus not high at-risk for externalizing behaviors.

Table 11

Descriptive Statistics for Student Gains in Treatment Group for Externalizing Behavior

	High At-Risk		Not High At-Risk	
Measure	*N*	Mean (SD)	*N*	Mean (SD)
Pennies Test	6	2.00 (4.56)	71	2.83 (3.29)
Texas Word Problems	6	0.33 (2.34)	71	3.32 (3.48)
Open Equations	6	1.17 (5.85)	71	5.54 (5.38)

Table 12

ANOVA Results for Student Gains in Treatment Groups for Externalizing Behavior

	F	*p*-value	Cohen's *d*
Pennies Test	.033	.57	0.21
TWP	4.24	*.04	1.01
Open Equations	3.61	.06	0.78

Note. * = statistically significant at p <.05

Table 13

Descriptive Statistics for Student Gains in Control Group for Externalizing Behavior

	High At-Risk		Not High At-Risk	
Measure	*N*	*M* (*SD*)	*N*	*M* (*SD*)
Pennies Test	6	1.83 (3.49)	44	2.46 (3.20)
Texas Word Problems	6	-0.67 (0.82)	44	0.57 (2.53)
Open Equations	6	1.83 (4.31)	44	3.91 (4.74)

A regression model was used to determine if higher externalizing scores predicted Texas Word Problem gain scores. In this model, continuous externalizing scores were used and not student at-risk status, where higher scores indicate more at-risk. Results of the regression model determined that as the externalizing scores increases by 1 point, students will perform 0.21 points lower on Texas Word Problems. A weak correlation (r = .205, p = 0.07) existed between the two variables. Table 14 shows the regression model statistics.

Table 14

Regression Model Statistics for Word-Problem Gains by Externalizing Scores

Model	*B*	Mean (SD)	β
(Constant)	3.85	0.57	
Externalizing Score	-0.21	0.12	-0.21

Note. R^2 = 0.04 (p = .07).

Internalizing. I ran similar analyses for students who were high at-risk versus not high at-risk for internalizing behaviors within the treatment group. The results, however, were not statistically significant for any outcome. A regression model was also conducted to determine predictability for internalizing behaviors on Texas Word Problem gains, however, that score was not statistically significant (R^2= 0.03, F(1, 75) = 2.07, p = .15). In other words, students with higher internalizing scores performed lower on the Texas Word Problems, however, those scores were not statistically significant.

Summary of Results

The results of the current study did confirm the hypothesis that students with MD exhibit higher rates of externalizing and internalizing behaviors than students without MD, with medium effect sizes noted. With regard to the second research question, students with MD in the treatment groups outperformed their peers in the control group on Texas Word Problems, noting a large effect size. For the third research question, the hypothesis was confirmed that students that exhibited high at-risk externalizing and internalizing behaviors performed lower, in both the treatment and control groups, on all outcomes. Marked differences existed in the externalizing treatment group specifically with the Texas Word problems, however, those results were not statistically significant and did not have strong correlations.

Chapter V: Discussion

The purpose of this study was threefold. First, the present study sought to examine the behavioral profiles of students with MD and without MD to determine if similar behavioral profiles existed between the groups. Specifically, did students with MD demonstrate higher scores in externalizing and internalizing behaviors on the SRSS? Second, this study sought to examine the efficacy of a word-problem intervention on students with MD compared to students in the control group. Last, this study examined the effect of behavior on the efficacy of a word-problem intervention.

Results indicate that students with MD display higher rates of externalizing and internalizing behaviors compared to students without MD. With regard to the second research question, students in the treatment groups outperformed students in the control group on all three outcome measures, however, only the Texas Word Problems were statistically significant. The third research question revealed that students that are observed to be high at-risk for externalizing and internalizing behaviors according to the SRSS in the treatment groups did not display the same pattern of gains on the outcome measures as their peers who were not determined to be high at-risk.

Research Question 1

Students with MD had statistically significantly higher externalizing and internalizing behaviors on the SRSS as rated by their classroom teachers versus their peers without MD. Previous research suggests that 42% to 93% of students that demonstrate challenging behaviors have an evidence of an MD (Epstein et al., 2005).

Therefore, it was expected that students with MD would demonstrate higher rates of at-risk behaviors.

Research Question 2

In the second research question, students receiving the word-problem intervention outperformed students in the control group on every outcome measure. The only statistically significant outcome, however, was the Texas Word Problems showing a large effect size. Texas Word Problems consisted of seven double-digit word problems. Students were given one point for the answer and one point for correctly labeling the word problem. Texas Word Problems closely mirrored the Pirate Math intervention in that, students were specifically asked three Total schema problems, three Difference schema problems, and one Change schema problem. An additional analysis was conducted to examine if differences in posttest gains were significant between the two treatment groups in the parent study (e.g., Pirate Math with equal sign instruction and Pirate Math without equal sign instruction). The difference between the two groups on posttest gains, however, was not statistically significant. This nonsignificant result allowed for a collapse of the two treatment groups into one for this analysis.

Compared to the Texas Word Problems test, the Pennies Test consisted of 14 single-digit mathematics word problems, where students solved only two Total and six Difference word problems. The rest of the Pennies Test (i.e. six word-problems) were Change word problems. By the timing of the present study's posttest, students in the treatment groups had not received explicit instruction on Change schemas. Therefore, on the Pennies Test, students in the treatment and control groups did not vary significantly.

The lack of statistical significance could be attributed to the heavy emphasis on Change problems within the Pennies Test.

The Open Equations test was a 6-min timed test consisting of 30 standard and non-standard equations (i.e. 2 + 4 = __+ 3). Students in the treatment group did outperform students in the control group in Open Equations but, again, it was not statistically significant. The Open Equations test is intended to test the effects between the two treatment groups (i.e. Pirate Math with equal sign instruction and Pirate Math without equal sign instruction). Students in the Pirate Math with equal sign instruction group receive explicit instruction in solving standard and nonstandard equations. Therefore, subsequent analysis was conducted to examine the difference in the two groups and while the Pirate Math with equal sign instruction group did score higher on the Open Equations, it was not statistically significant. It could, however, account for the difference noted between the treatment groups and the control group at posttest.

Research Question 3

My third research question investigated the efficacy of the word-problem intervention based on student behavioral profiles. That is, do higher externalizing or internalizing behaviors predict student response to the word-problem intervention? Two different analysis were used to attempt to answer this question for both internalizing and externalizing behavior profiles.

To answer this question for externalizing behaviors, I compared students whose behavioral scores were between 9 and 21 on the SRSS, indicating a high at-risk status, to those students who scored below 9 and their gains on the three outcome measures.

Students without high at-risk status (*n* = 71) outperformed students who were high at-risk for externalizing behaviors (*n* = 6). Most notably, the test of Texas Word Problems displayed the greatest differences, where students that were not high at-risk significantly outperformed students deemed high at-risk with a Cohen's *d* of 1.01 indicating large effect size. Students with high at-risk for internalizing behaviors (*n* = 15) did not perform significantly different than students not high at-risk (*n*= 62). Interestingly, the control groups for externalizing and internalizing behaviors did not display any statistical significance. This confirms the hypothesis that students with externalizing types of challenging behaviors do not respond to academic interventions as well as students without challenging behaviors and in turn, may need more intensive intervention to see higher effects.

A regression model was then used to determine if higher externalizing scores, therefore, predicted gains on the Texas Word Problem test. As externalizing behaviors increase, student's word-problem gains decrease by 0.21. The regression model, however, yielded marginally statistical significance with a *p*-value of 0.07. This is somewhat surprising but may be because of the low numbers of students (*n* = 6) with high at-risk scores in this sample.

Implications

Results from this study have several implications. First, this study was an exploratory study to determine if there was (a) a link between students with MD and behavior and (b) if there was a link, was behavior impacting student posttest gains in the word-problem intervention? This study did establish a link between students with MD and behavior with this sample population. Students with MD displayed higher externalizing and internalizing scores over their peers without MD. This further confirms the findings from Epstein and colleagues (2012) suggesting that students with challenging behaviors also exhibit challenges with mathematics achievement.

Further implications exist for mathematics researchers when designing word problem interventions in kindergarten through grade 12. While there were not several students in third grade that exhibited high at-risk externalizing behaviors, research indicates that the rates of EBD identification increase as students approach middle and high school (Nelson et al, 2004). Additionally, the mathematics achievement gap also widens as students with challenging behaviors progress through school (Wei et al., 2012). Anecdotally, between screening and posttest, several students (n = 7) that scored high on the SRSS for externalizing behaviors were lost due to moving schools, long term suspension, or extreme behaviors during pre-test and therefore not able to complete testing. While this is not surprising due to the somewhat behavioral-related transiency of students with challenging behaviors, it is a consideration when examining the impact of behavior on academics. Therefore, those intervention researchers should continue to be aware of the co-morbidity that exists between behavior and mathematics and design word-problem interventions accordingly.

This study also establishes the need to intensify interventions for students that have co-occurring MD and high at-risk externalizing behaviors. While higher numbers of students with MD had at-risk internalizing behaviors (i.e., n = 15 versus n = 6), their posttest gains did not appear to be significantly impacted according to the outcome measures. Students with high at-risk externalizing behaviors, however, were significantly impacted compared to students not high at-risk for externalizing behaviors. Students that had higher at-risk performed worse on all outcome measures than did their peers without high at-risk. Notably, those students performed significantly lower on the Texas Word problems. Therefore, this establishes the need to intensify the word-problem intervention to ensure all students can make adequate gains. In a more broader sense, students with challenging behaviors, specifically students with EBD, demonstrate greater achievement gaps as they progress through their academic careers (Wei et al., 2012). Therefore, intensifying interventions in elementary interventions is increasingly important.

Limitations and Recommendations for Future Research

There are several limitations in the present study. First, the researcher was not able to obtain behavioral data on 25% of the consented students for various reasons. This, therefore, could have impacted the results of the findings. Additionally, the SRSS is primarily utilized as a behavioral screener in MTSS and intended to be given three times a year to monitor student behavioral progress on a macro scale. In the present study, the SRSS was only given once prior to the beginning of the word-problem intervention. Therefore, students identified as high at-risk for externalizing and internalizing behaviors at the beginning may not necessarily be the same students at the end of the present study. Also, in the present study, very few students were identified as having high externalizing (n = 6) and internalizing behaviors (n =15). That may have impacted the results of the data due to the small number of students in those groups. Lastly, the present study evaluated students after they had completed Lesson 30 in Pirate Math. Results may have been different if students had completed all of the lessons (i.e., Lesson 48) in the intervention series. The decision to not finish out the intervention lessons was made primarily due to timing constraints.

Future Research

Future research should continue to explore the link between behavior and mathematics. First, research should be conducted in other areas of mathematics intervention research to identify if behavior impacts more than just word-problem solving. Additionally, future research should explore how to intensify word-problem interventions for students with at-risk externalizing behaviors. This intervention was conducted one-on-one and already considered a form of intensification (National Center on Intensive Intervention, 2016). Therefore, future research should continue to examine the results of (a) adjusting the Pirate Math intervention components, specifically increasing reinforcement rates or performing preference assessments to determine specific reinforcers or (b) consider increasing dosage, utilize more three-dimensional representations, explicitly teach transfer skills to increase opportunities for practice, or other recommendations from the National Center on Intensive Intervention (2016). Last, future research should replicate this study in geographically diverse areas to examine if similar results exist.

Summary

In conclusion, the purpose of this study was to examine the impact of behavior and mathematics, specifically the impact of behavior on a word-problem intervention. The intervention was delivered for 15 weeks by trained interventionists in a one-on-one setting. There was a statistically significant difference in the behavioral profiles of students with MD and those without MD. Additionally, students with MD in the treatment group outperformed students in the control group on all post-test outcomes, especially in the Texas Word Problems. Moreover, scores on the SRSS did not predict performance on the Texas Word Problems test but students with high at-risk externalizing behaviors did perform significantly lower on that outcome measure than their peers without high at-risk status. Students with co-occurring behavior and mathematics difficulties may need more intensive intervention to see similar results as their peers without behavioral challenges on a word-problem intervention.

Appendix

Appendix A. Student Risk Screening Scale (SRSS)

Student Risk Screening Scale - Internalizing and Externalizing (SRSS-IE) 2.0
ELEMENTARY USE

DATE:

TEACHER NAME:

0 = Never
1 = Occasionally
2 = Sometimes
3 = Frequently
Use the above scale to rate each item for each student.

Student Name	Student ID	Count	Steal	Lie, Cheat, Sneak	Behavior Problem	Peer Rejection	Low Academic Achievement	Negative Attitude	Aggressive Behavior	Emotionally Flat	Shy; Withdrawn	Sad; Depressed	Anxious	Lonely	SRSS TOTAL	SRSS-I5 TOTAL	SRSS-IE TOTAL
Example: Smith, Sally	*11111*	0	*0*	*0*	*3*	*1*	*3*	*3*	*3*	*2*	*2*	*2*	*3*	*0*	13	9	22
		1													0	0	0
		2													0	0	0
		3													0	0	0
		4													0	0	0
		5													0	0	0
		6													0	0	0
		7													0	0	0
		8													0	0	0
		9													0	0	0
		10													0	0	0
		11													0	0	0
		12													0	0	0
		13													0	0	0
		14													0	0	0
		15													0	0	0
		16													0	0	0
		17													0	0	0
		18													0	0	0
		19													0	0	0
		20													0	0	0
		21													0	0	0
		22													0	0	0
		23													0	0	0
		24													0	0	0
		25													0	0	0
		26													0	0	0
		27													0	0	0
		28													0	0	0
		29													0	0	0
		30													0	0	0

Appendix B. Pretest and Post-test

FIRST:_________________________ **LAST:**____________________________

PT ☐ **TWP** ☐ **OE** ☐

_____ 1. Alex has 8 pennies. Kris has 6 pennies. How many pennies does Alex need to give away to have as many as Kris?

_____ 2. Sue had 5 pennies. Then Mike gave her 2 more pennies. How many pennies does Sue have now?

_____ 3. Chelsea has 6 pennies. Max has 4 pennies. How many pennies does Max have less than Chelsea?

_____ 4. Nina had 9 pennies. Then she gave 3 pennies to Anthony. How many pennies does Nina have now?

_____ 5. Janet has 3 pennies. Andy has 5 more pennies than Janet. How many pennies does Andy have?

_____ 6. Carol had 4 pennies. Then Nick gave her some more pennies. Now Carol has 6 pennies. How many pennies did Nick give her?

_____ 7. Claire has 4 pennies. Ben has 9 pennies. How many more pennies does Claire need to have as many as Ben?

______ 8. Jen had 7 pennies. Then she gave some pennies to Joe. Now Jen has 2 pennies. How many pennies did she give to Joe?

______ 9. Emily has 3 pennies. John has 6 pennies. How many pennies do they have altogether?

______ 10. Maria and Kevin have 8 pennies together. Maria has 3 pennies. How many pennies does Kevin have?

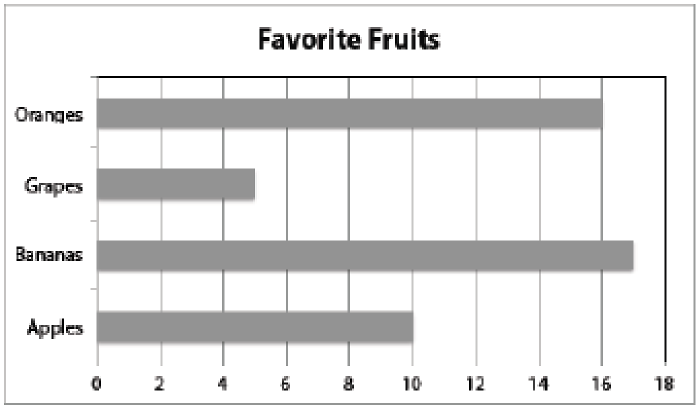
______ 11. Ashley has 7 pennies. Jason has 4 pennies less than Ashley. How many pennies does Jason have?

______ 12. Dennis has 7 pennies. Molly has 5 pennies. How many pennies does Dennis have more than Molly?

______ 13. Karen had some pennies. Then Matt gave her 4 more pennies. Now Karen has 6 pennies. How many pennies did she have to start with?

______ 14. Lisa had some pennies. Then she gave 3 pennies to Bill. Now Lisa has 5 pennies. How many pennies did Lisa have to start with?

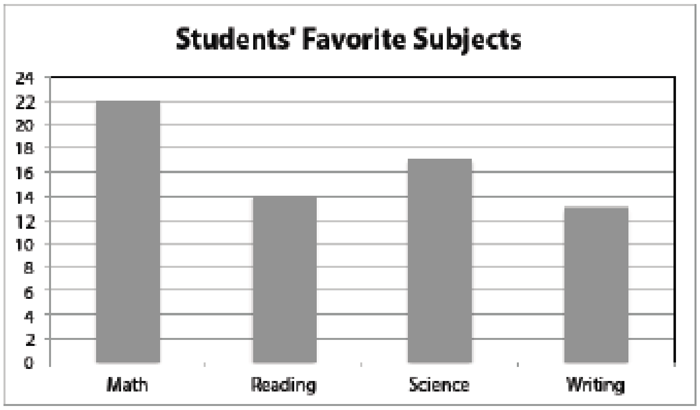
A.



The graph shows the favorite fruit of third-grade students. How many students chose oranges or bananas?

B. Abby is 5 years older than her brother. If Abby is 16 years old, how old is her brother?

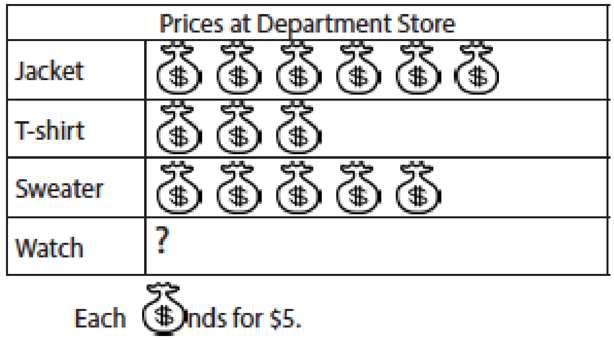
C.



The graph shows the favorite subject of third-grade students. How many more students chose Math than chose Writing?

D. Jack brought 78 cookies to school on his birthday. Then, he gave some cookies to his friends. Now, Jack has 23 cookies. How many cookies did Jack give away?

E.



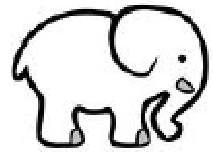
The graph shows the prices of items at the department store. A new t-shirt and watch cost $41 in all. How much does a watch cost?

F. Last year, there were 11 trumpet players in the band. This year, 14 new trumpet players and 4 tuba players joined the band. How many trumpet players are in the band now?

G. Charlie has 24 more baseball cards than football cards. He has 45 football cards. How many baseball cards does Charlie have?

H. Donna and Natasha folded 96 paper cranes. Donna folded 25 paper cranes. How many paper cranes did Natasha fold?

Open Equations



1. ____ + 3 = 7
2. 2 = 7 – ____
3. ____ = 4
4. 6 = 2 + ____
5. ____ – 4 = 3
6. 3 + 5 = 4 + ____
7. ____ = 7 – 4
8. ____ – 6 = 2
9. 9 = ____ + 4
10. 8 – 6 = ____ – 3
11. ____ – 3 = 8 – 2
12. 5 = ____ + 3
13. 5 = 9 – ____
14. 3 + ____ = 8

15. 5 + 4 = ____ + 2

16. 9 – ____ = 6

17. 7 + 2 = ____

18. ____ + 4 = 5 + 2

19. 7 = ____ – 2

20. 7 – ____ = 5

21. 5 + ____ = 9

22. 3 + ____ = 2 + 7

23. 6 = ____ – 2

24. 9 – 6 = 7 – ____

25. ____ + 6 = 9

26. 7 = ____

27. ____ = 2 + 6

28. 8 – 3 = ____

29. 6 – ____ = 7 – 3

30. 7 = 4 + ____

Appendix C: Sample Fidelity Sheet

Fidelity Checklist
Lesson 17 for PMES

Tutor:
Teacher:
Reliability completed by:
School:
Student:
Percentage:

MATH FACT FLASH CARDS
____1. Tutor holds up flash cards and explains that they have addition and subtraction problems
____2. Tutor explains that the student will have 1 min to solve as many problems as he/she can
____3. Student says as many flash cards as he/she can in 1 min
____4. Tutor and student count the cards in the pile
____5. Tutor asks student if he/she can beat score
____6. Tutor uses the same flash cards and student says as many as he/she can in 1 min
____7. Tutor and student graph the high score for the day on the graph
EQUATION QUEST
____8. Tutor asks student what the equal sign means and reviews that it means the same as
____9. Tutor asks student what is means to isolate the X
____10. Tutor introduces problem (A) asks student to read the number sentence 10-6=4
____11. Tutor and student discuss the minus sign
____12. Tutor reviews with student that the equal sign means the same as
____13. Tutor introduces problem (B) and reads the number sentence 5-2 is the same as blank
____14. Tutor and student writes X in the blank
____15. Tutor place 5 cubes on one side and subtracts 2 cubes
____16. Tutor place X on the other side of the equal sign
____17. Tutor reviews isolating X with the student
____18. Tutor and student determine X is the same as 3
____19. Tutor reviews with student that the equal sign means the same as
____20. Tutor introduces problem (C) and reads the number sentence
____21. Tutor and student writes X in the blank
____22. Tutor place X on one side of the equal sign
____23. Tutor place 8 cubes on one side and subtracts 3 cubes
____24. Tutor and student determine X is the same as 5
____25. Tutor and student reviews 5 is the same as 8 minus 3
BUCCANEER PROBLEMS
____26. Tutor displays Buccaneer Problems
____27. Tutor reviews that in a Total problem, two or more parts are put together into a total
____28. Tutor and student review the Total equation of P1+P2=T
____29. Tutor asks student what two things they need to include in their answer (number and label)
____30. Tutor asks student what a label is and they review that a label tells us about our missing informa

References

Alter, P. (2012). Helping students with emotional and behavioral disorders solve mathematics word problems. *Preventing School Failure*, *56*, 55−64. doi:10.1080/1045988X.2011.565283

Alter, P. Brown, E.T., & Pyle, J. (2011). A strategy−based intervention to improve math word problem−solving skills of students with emotional and behavioral disorders. *Education and Treatment of Children, 34*, 535−550. doi:10.1353.2011.0028

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: Author

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976-82. *Ambulatory Pediatrics, 5,* 281–289. doi:10.1367/A04-209R.1

Blackorby, J., Wagner, M., Cameto, R., Davies, E., Levine, P., Newman, L., . . . Sumi, C. (with Chorost, M., Garza, N., & Guzman, A. M.). (2004). *Engagement, academics, social adjustment, and independence: The achievements of elementary and middle school students with disabilities*. Menlo Park, CA: SRI International.

Bullis, M. & Yovanoff, P. (2006). Community employment experiences of formerly incarcerated youth. *Journal of Emotional and Behavioral Disorders*, *14*, 71−85.

Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K. L. S., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disability Quarterly, 31*, 11–20.

Cook, B., Buysse, V., Klingner, J., Landrum, T. McWilliam, R., Tankersley, M & Test, D. (2015). *Council for exceptional children standards for evidence-based practices in special education*. Arlington, VA: Council for Exceptional Children.

Cutler, C. S. & Monroe, E. E. (2006). Sixth graders' oral retellings of compare word problems. *Focus on Learning Problems in Mathematics, 28*(2), 309–313.

Driver, M. K., & Powell, S. R. (2015). Symbolic and nonsymbolic equivalence tasks: The influence of symbols on students with mathematics difficulty. *Learning Disabilities Research and Practice, 30,* 127–134. doi:10.1111/ldrp.12059

Drummond, T. (1994), The Student Risk Screening Scale (SRSS). Retrieved from www.CI3T.org.

Epstein, M. H., Nelson, J. R., Trout, A. L., & Mooney, P. (2005). Achievement and emotional disturbance: Academic status and intervention research. In M. H. Epstein, K., Kutash, K., & Duchnowski, A.J. (Eds.), *Outcomes for children and youth with emotional and behavioral disorders and their families: Programs and evaluation practices* (2nd ed., pp. 451-477). Austin, TX: PRO-ED.

Epstein, K. Forness, S. R., Freeman, S. F. N., Paparella, T., Kauffman, J. M., & Walker, H. M. (2012). Special education implications of point and cumulative prevalence for children with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *20*, 4−18.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., Karns, K., & Dutka, S. (1997). Enhancing students' helping behavior during peer tutoring with conceptual mathematical explanations. *The Elementary School Journal*, *97*, 223−250.

doi:10.1086/461863

Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first−grade children's mathematical development with peer−assisted learning strategies. *School Psychology Review, 31*, 569−583

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences,* 20, 89−100. doi:10.1016/j.lindif.2009.09.003

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D,...Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101,* 561−576. doi:10.1037/a0014701

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third−grade students with math and reading difficulties. *Exceptional Children, 74,* 155−173.

Geary, D.C. (2011). Cognitive predictors of achievement growth in mathematics: A five year longitudinal study, *Developing Psychology*, *47*, 1539–1552. doi:10.1037/a0025510

Geary, D.C., Hoard, M.K., & Bailey, D.H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning*

Disabilities, 45, 291–307. doi:10.1177/0022219410392046

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability, *Child Development, 78,* 1343–1359. doi:10.1111/j.1467-8624.2007.01069.x

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Griffin, C. & Jitendra, A. (2009). Word problem-solving instruction in inclusive third-grade mathematics classrooms. *Educational Psychology, 102*, 187–201. doi: https://doi-org.ezproxy.lib.utexas.edu/10.3200/JOER.102.3.187-202

Hedges, L. V. (2013). Recommendations for practice: Justifying claims of generalizability. *Educational Psychology Review, 25*(3), 331–337. doi: 10.1007/s10648-013-9239-x

Higgins, J. W., Williams, R. L., & McLaughlin, T. F. (2001). The effects of a token economy employing instructional consequences for a third-grade student with learning disabilities: A data-based case study. *Education and Treatment of Children, 24*, 99–106.

Hodge, J., Riccomini, P.J., Buford, R., & Herbst, M. (2006). A review of instructional interventions in mathematics for students with emotional and behavioral disorders. *Behavioral Disorders, 31,* 297−311.

Horner, R. H., Carr, E. G., McGee, G., Odom, S. & Wolery, M. (2005) The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165–179. doi: 10.1177/001440290507100203

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Institute of Education Sciences. (2014). What Works Clearinghouse procedures and standards handbook (Version 3.0). Washington, D.C.: Author. Retrieved from http://www.ies.ed.gov.

Jordan, N. C. & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33,* 567–578. doi:10.1177/002221940003300605

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case design technical documentation. Retrieved from *What Works Clearinghouse* website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Lane, K. L. (2004). Academic instruction and tutoring interventions for students with emotional/behavioral disorders: 1990 to present. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds), *Handbook of research in emotional and behavioral disorders* (pp. 462−486). New York, NY: Guilford Press.

Lane, K. L. L., Barton-Arwood, S. M., Nelson, J. R., & Wehby, J. (2008). Academic performance of students with emotional and behavioral disorders served in a selfcontained setting, *Journal of Behavioral Education, 17*, 43−62. doi:10.1007/s10864-007-9050-1

Lane, K. L., Little, M. A., Casey, A. M., Lambert, W., Wehby, J. H., Weisenbach, J. L., & Phillips, A. (2009). A comparison of systematic screening tools for emotional and behavioral disorders: How do they compare? *Journal of Emotional and Behavioral Disorders, 17,* 93−105. doi: 10.1177/1063426608326203

Lane, K. L., Oakes, W. P., Swogger, E. D., Schatschneider, C., Menzies, H., M., & Sanchez, J. (2015). Student risk screening scale for internalizing and externalizing behaviors: Preliminary cut scores to support data-informed decision making. *Behavioral Disorders, 40,* 159–170.

Lane, K. L., Parks, R. J., Kalberg, J. R., & Carter, E. W. (2007). Systematic screening at the middle school level: Score reliability and validity of the student risk screening scale. *Journal of Emotional and Behavioral Disorders, 15,* 209−222. doi: 10.1177/10634266070150040301

Lee, Y. Y., Sugai, C., & Horner, R. H. (1999). Using an instructional intervention to reduce problem and off-task behaviors. *Journal of Positive Behavior Interventions, 1*, 195–204. doi:10/1177/109830079900100402

Maggin, D. M., Wehby, J. H., Farmer, T. W., & Brooks, D. S. (2016). Intensive interventions for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 24*(3), 127–137. doi: 10.1177/1063426616661498

Matthews, P. G. & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing selfexplanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, *104*, 1–21. doi:10.1016/j.jecp.2008.08.004

Merrell, K. W. & Walker, H. M. (2004). Deconstructing a definition: Social maladjustment versus emotional disturbance and moving the EBD field forward. *Psychology in the Schools, 41*, 899–910. doi:10.1002/pits.20046

Miller, S. P., & Hudson, P. J. (2006). Helping students with disabilities understand what mathematics means. *Teaching Exceptional Children, 39*, 28–35. doi: 10.1177/004005990603900105

Mooney, P., Epstein, M. H., Reid, R., & Nelson, J. R. (2003). Status and trends of academic intervention research for students with emotional disturbance. *Remedial and Special Education, 24*, 272−287. doi:10.1177/07419325030240050301

Mulcahy, C. A. & Krezmien, M. P (2009). Effects of a contextualized instructional package on the mathematics performance of secondary students with EBD. *Behavioral Disorders, 34*, 136−150.

Mulcahy, C. A., Krezmien, M. P., & Travers, J. (2016). Improving mathematics performance among secondary students with EBD: A methodological review. *Remedial and Special Education, 37*, 113−128. doi:10.177/0741932515579275

National Center on Intensive Intervention. (2016). *How to intensify academic Interventions.* Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Intensive Intervention. Retrieved from http://www.intensiveintervention.org.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report* *of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.

Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). An investigation of the academic achievement of K−12 students with emotional and behavioral disorders in public school settings. *Exceptional Children, 71*, 59–73. doi: 10.1177/001440290407100104

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. *Psychology Science, 25*(4), 903–-910. doi:10.1177/0956797613518351

Oakes, W. P. Wilder, K., Lane, K. L., Powers, L., Yokoyama, L., O'Hare, M. E., & Jenkins, A. B. (2010). Psychometric properties of the Student Risk Screening Scale: An effective tool for use in diverse urban elementary schools. *Assessment for Effective Intervention, 35,* 231−239. doi: 10.1177/1534508410379796

Peltier, C., & Vannest, K. J. (2016). Schema-based strategy instruction and the mathematical problem-solving performance of two students with emotional and behavioral disorders. *Education and Treatment of Children, 39*, 521−544. doi: 10.1353/etc.2016.0023

Powell, S.R. (2007). Open Equations. Available from S. R. Powell, 1912 Speedway Stop 5300, Austin, TX, 78712.

Powell, S.R. (2011). Solving word problems using schemas: A review of the literature. *Learning Disabilities Research and Practice. 26,* 94–108. doi:10.111/j.1540-5826.2011.00329.x

Powell, S. R. (2012). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal, 112*, 627–648. doi:10.1111/j.1540-5826.2011.00329.x

Powell, S. R. (2015). The influence of symbols and equations on understanding mathematical equivalence. *Intervention in School and Clinic*, *50*, 266–272. doi:10.1177/0731948714564574

Powell, S.R. (2016). An analysis of mathematics vocabulary from kindergarten through eighth grade. Manuscript submitted for publication.

Powell, S. R., & Barnes, M.A. (2015-2019). *Developing connections between word problems and mathematical equations to promote word-problem performance among students with mathematics difficulty.* Institute of Education Sciences, Goal 3, R324A150078 ($3,013,726).

Ralston, N. C., Benner, G. J., Tsai, S. F., Riccomini, P. C., & Nelson, J. R. (2014). A best evidence synthesis of the effects of mathematics instruction for students with emotional and behavioral disorders. *Preventing School Failure, 58*, 1−16. doi: 10.1080/1045988X.2012.726287

Reid, R., Gonzalez, J.E., Nordness, P.D., Trout, A., & Epstein, M.H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38*, 130–144. doi: 10.1177/00224669040380030101

Riley, M. S & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities of solving problems. *Cognition and Instruction*, *5*, 49–101.

doi:10.1207/s1532690xci0501_2

Rules and Regulations. (2006). Federal Register. *71*(156), 46756.

Ryan, J. B., Pierce, C. D., and Mooney, P. (2008). Evidence-based teaching strategies for students with EBD. *Beyond Behavior*, 17(*3*), 22−29.

Templeton, T. N., Neel, R. S., & Blood, E. (2008). Meta-analysis of math interventions for students with emotional and behavioral disorders. *Behavioral Disorders, 16,* 226−239. doi: 10.1177/1063426608321691

Tournaki, N. (2003). The differential effects of teaching addition through strategy instruction versus drill and practice to students with and without learning dis-abilities. *Journal of Learning Disabilities, 36*, 449–458. doi: 10.1177/00222194030360050601

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2015 Mathematics Assessment.

U.S. Department of Education, National Center for Education Statistics. (2016). *Digest of Education Statistics, 2014* (NCES 2016−006), Chapter 2.

U.S. Department of Education, Office of Special Education Programs, Data Analysis System. (2001). *Percentage of students with disabilities identified as having emotional disturbance.* Retrieved from http://cecp.air.org/resources/20th/eligchar.asp

Wagner, M., Newman, L., & Cameto, R. (2004). *Changes over time in the secondary school experiences of students with disabilities. A report of findings from the*

National Longitudinal Transition Study (NLTS) and the National Longitudinal Transition Study−2 (NLTS2). Menlo Park, CA: SRI International.

Wehmeyer, M. L., Sands, D. J., Doll, B., & Palmer, S. (1997). The development of self-determination and implications for educational interventions with students with disabilities. *International Journal of Disability, Development and Education*, *44*, 305–32. doi: 10.1080/0156655970440403

Wei, X., Lenz, K. B., & Blackorby, J. (2012). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education*, *34*, 154−165. doi:10.1177/0741932512448253

Willis, G. B. & Fuson, K. C. (1988). Teaching children to use schematic drawings to solve addition and subtraction word problems. *Journal of Educational Psychology*, *80*, 192–201. doi: 10.1037/0022-0663.80.2.192

Xin, Y.P . & Zhang D. (2009). Exploring a conceptual model-based approach to teaching situated word problems. *The Journal of Educational Research*, *102*, 427–441. doi: 10.1080/00220671.2011.593983