The Dissertation Committee for Gavin Walter Watts Certifies that this is the approved version of the following dissertation:

Effects of Cross-Age Tutors with EBD on the Mathematics Performance of At-risk Kindergarteners

Committee:

Diane Pedrotty Bryant, Supervisor

Nathan H. Clemens

Terry S. Falcomata

Kathleen H. Pfannenstiel

Effects of Cross-Age Tutors with EBD on the Mathematics Performance of At-risk Kindergarteners

by

Gavin Walter Watts

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

First and foremost, I must thank my parents, Geoff and Leslie, for whom I would not be here, literally. They are the greatest teachers, role models, and cheerleaders that I could ever ask for. This achievement would not have been possible without their steadfast support, and their positive influence continues to propel me forward in my life and career. I'd also like to thank my brother Mitchell for continuing to inspire me through his own high aspirations and achievements. His courage has been exhibited through his service to his country, and now, his local community. I am also motivated by his creativity and constant drive to undertake new challenges and projects. Next, I'd like to thank Lauren Grant for her love, understanding, and help in balancing all of the responsibilities that I have taken on over the past four years. She is an absolute support, a calming influence, and a true compass, always pointing me in the direction of what I truly want and need… which is usually a trip to Marfa, TX. I also have to thank Joel Kerr, who has served as my steadfast collaborator and sounding board for professional and creative endeavors over the past decade. I admire and appreciate Joel's ability, and willingness to put in his own time and effort in order to produce ideas that elevate any project. He always matches and expounds upon my drive and aspirations to create and develop something meaningful. The completion of this project would not have been possible without his involvement and feedback. Additionally, to all of my extended family and friends who have shown an interest in my career and development as an educator, I am forever grateful, and hope that I have made you proud. Finally, I must thank Dr. Diane Bryant who has served as my advisor, mentor, collaborator, and now, colleague. She has provided me with a robust amount of training, experience, and feedback that has been essential to my development towards becoming an expert in my field, and henceforth, the obtainment of this doctoral degree.

Thank you all.

Effects of Cross-Age Tutors with EBD on the Mathematics Performance of At-risk Kindergarteners

Gavin Walter Watts, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Diane Pedrotty Bryant

Challenges with numerical proficiency at an early age can lead to substantial gaps in learning and are associated with detrimental long-term outcomes. Additionally, the academic and behavioral needs of students with emotional-behavioral disorders (EBD) have been identified as some of the most challenging to address. The purpose of this study was to identify the effects and related outcomes of utilizing cross-age tutors (i.e., older students) with, or at-risk for EBD to deliver a number line board game intervention to kindergarten students at-risk for mathematics disabilities. A concurrent multiple baseline design across participants was utilized to evaluate results related to the following research questions: (1) What effects will a number line game delivered by a cross-age tutor with EBD have on the early numeracy knowledge and skills of kindergarten students at-risk for math disabilities? (2) Can students with EBD effectively serve in the role of cross-age tutors (i.e., implement instruction with fidelity and increase tutees' number sense skills)? (3) What effects will the training and implementation of the cross-age tutoring program have on the tutors' behavioral performance as well as overall risk status for EBD? Tutoring sessions took place for 25–30 minutes, three times per week, over 10 weeks. Results suggest this cross-age tutoring program to be an effective and feasible model for significantly improving mathematical performance of tutees at-risk for mathematics disabilities and, to a lesser extent, the behavioral ratings of students with EBD. Distal measures showed the intervention's moderate effect on tutees' mathematics performance and large effect on decreasing tutors' risk-status for EBD. Tutors implemented the intervention procedures with high rates of fidelity and, in combination with the significant gains by their tutees, demonstrated the ability of students with EBD to effectively serve as cross-age tutors. In assessing the social validity of this instructional model, the implementing special educator rated the intervention to be effective and beneficial, although challenges were identified in the area of scheduling. All tutors and tutees perceived the program as effective in promoting mathematics skills for the tutees and positive behavioral developments for the tutors. Limitations, implications for practice, and areas of future research are discussed.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

CHILDREN WITH MATHEMATICS DIFFICULTIES AND EARLY NUMERACY

Young children with or at-risk for mathematics learning disabilities (LD) are often challenged with basic number sense knowledge and skills (Geary et al., 2009; Mazzocco, Feigenson, & Halberda, 2011). Students with mathematics LD struggle to develop adequate number sense knowledge and skills required to facilitate later fact and computation skills (Locuniak & Jordan, 2008). Students who are typically developing in early numeracy knowledge and skills usually achieve high number sense skills by first-grade, while students with mathematics LD have been shown to have continuing deficits in these same skills through third-grade (Desoete & Grégoire, 2006).

Research suggests between 5% to 10% of school-age children are diagnosed with mathematics LD (Fuchs, Fuchs, & Hollenbeck, 2007). The National Assessment of Educational Progress (NAEP; 2015) mathematics scores indicate that students with disabilities are struggling to obtain basic mathematics skills. For example, results show that students with disabilities in fourth-grade are three times more likely to score below the basic level, and that the overall population of students with disabilities, who scored below the *basic level*, has increased from 43% in 2005 to 45% in 2015 (NAEP; 2015). Additionally, low socioeconomic status appears to be a key factor related to early struggles in mathematical development (Jordan, Kaplan, Ramineni, & Locuniak, 2009).

Mathematics difficulties can be identified in the early grades and can lead to difficulties that are more serious, if left untreated. For example, a range of differences in early numeracy knowledge and skills can be identified as early as age five (Aunio, Hautamäki, Sajaniemi, & Van Luit, 2009; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004). Furthermore, these early discrepancies in number competence have been shown to have significant short-term and long-term outcomes for students with mathematics LD (Jordan, Glutting, & Ramineni, 2010; Jordan, Kaplan, Locuniak, & Ramineni, 2007; Kavkler, Aubrey, Tancig, & Magajna, 2000; Locuniak & Jordan, 2008).

In the short term, if students have been identified with a mathematics LD by age 7 or 8, it is likely that they were correctly identified as being at-risk for developing that disability in kindergarten (Stock, Desoete, & Roeyers, 2009; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011). Additionally, correlations have been identified between early numeracy knowledge/skills and mathematics achievement in the later elementary years (Gersten, Jordan, & Flojo, 2005; Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Mazzocco & Thompson, 2005; Morgan & Farkas, 2009).

Long term, students identified as having deficits in early numeracy knowledge and skills have shown continuing low performance on future measures of mathematics achievement (Baker, Gersten, & Lee, 2002; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Jordan, Kaplan, Locuniak & Ramineni, 2007). Children with low performance in early numeracy knowledge and skill areas may struggle to develop the conceptual foundations that will support the learning of more advanced mathematics (Van Luit & Schopman, 2000). Interestingly, when comparing the relation between early and later literacy skills and early and later mathematics achievement, mathematics proves to be a stronger predictor (Duncan et al., 2007). Morgan, Farkas, and Wu (2009) analyzed the Early Childhood Longitudinal Study data set and found that low mathematics achievement in the early elementary years has been related to poor performance in mathematics in later grades. Moreover, kindergarteners identified with low achievement on early numeracy measures were likely to have continuing struggles in mathematics in the late elementary and middle school years at a factor of 17 times that of their typically achieving peers (Morgan, Farkas, Hillemeier, & Maczuga, 2016).

These detrimental outcomes concerns have been shown in the Common Core State Standards initiative that includes raised expectations for early numeracy knowledge and skills (CCSS Initiative, 2010). In fact, promoting learning in the early years of a child's educational career can show positive benefits long-term (Clements & Sarama, 2011). This is especially true for kindergarten students from low-economic-status families who are at-risk for developing a disability (Baroody, Eiland, & Thompson, 2009; Dyson, Jordan, & Glutting, 2011). The National Mathematics Advisory Panel (NMAP, 2008) place a strong importance on early mathematics interventions, and preventative interventions, that have a strong foundation in whole number concept development and proficiency. Early intervention in these foundational concepts and skills is supported for students at risk for mathematics LD in an effort to deter long-term deficits and future barriers to learning advanced mathematical concepts (National Council of Teachers of Mathematics [NCTM], 2006; NMAP, 2008). If not effectively addressed, early difficulties in acquiring essential numeracy skills can persist into long-term challenges that may become impervious to intervention (Geary, 1993; Jordan, Kaplan, & Hanich, 2002).

One early numeracy intervention with a foundation in whole number concepts is a number line board game, which is based on theoretical frameworks and empirical research connections to the mental number line (Ansari, 2008; Hubbard, Piazza, Pinel, & Dehaene, 2005). Number knowledge is required to develop mental number lines, which are cognitive constructions of increasing number magnitude from left to right, horizontally. The connection between playing number line board games, and the development of number knowledge, and henceforth, the development of young children's' mental number line has been shown through the promotion of underlying early numeracy knowledge and skills in the areas of number magnitude, number line estimation, and number comparison (Ramani & Siegler, 2008; 2011; Siegler & Ramani, 2009; Whyte & Bull, 2008).

THEORETICAL FRAMEWORK

Number Competence for Early Numeracy and Number Line Board Games

The theoretical framework for early numeracy knowledge and skills can be situated in the work of researchers who have examined number competence and young children with mathematics difficulties. For example, Jordan, Kaplan, Ramineni, and Locuniak (2009) describe *number competencies* through three related domains: number, numerical relations, and arithmetic operations, all of which focus on early numeracy knowledge and abilities. Numbering requires the verbal counting sequence, knowledge of counting principles and cardinality (i.e., the ability to determine the total number of items in a set) through subitizing (i.e., immediate recognition of a quantity of a set) or by counting the individual items of the set. Counting principles contain the understanding that the total number is the final number counted when counting a set of items and that counting each item once is key. Related to the number relations competency is the concept of magnitude. Magnitude is a requisite feature of numeracy that supports number relations, which includes mathematic estimation and computation (Arnold, Fisher, Doctoroff, & Dobbs, 2002; Jordan, Kaplan, Olah, & Locuniak, 2006). Improving mental representations of magnitude, such as conceptual development of a mental number line in the early years improves students' abilities to develop more advanced mathematics skills later, such as basic arithmetic skills (e.g., addition and subtraction of whole numbers) (Jordan, Kaplan, Ramineni, & Locuniak, 2009). Research findings show that students with mathematics LD have difficulties in developing conceptual understanding of counting principles, and that these difficulties can affect the use of more advanced counting skills and the ability to solve arithmetic combinations (Geary, 2004; Griffin, 2004; Jordan, Kaplan, Ramineni, & Locuniak, 2009). Young children who exhibit underdeveloped counting strategies (e.g., counting on fingers) have shown difficulties in mastering arithmetic, which can manifest in later challenges in computational fluency (Gersten, Jordan, & Flojo, 2005; Jordan, Kaplan, & Hanich, 2002). In sum, *number competencies* are critical in developing advanced mathematical knowledge and skills.

As part of the theoretical framework for this study, research findings on the use of number line board games provide a rationale for employing this physical material and game in the early numeracy intervention. The use of physical materials, such as mathematics manipulatives and games in supporting mathematics instruction is often encouraged (Ainley, 1990; Ball, 1992), and in particular, number line board games have been shown to promote young students' knowledge of numerical magnitude and other early numeracy knowledge and skills such as counting, number line estimation, number identification, and arithmetic (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008). Research results on the use of number line board games indicate that number line estimation can support early numeracy development due to the requirement to assess magnitudes (Siegler & Booth, 2005).

Young children who have been identified in later grades as having mathematics disabilities have been shown to benefit from playing number line board games (Laski & Siegler, 2014; Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009). Children who engaged in number line board games showed significant developments in numerical competency in the areas of counting, numerical magnitudes, number comparison, and number line estimation, compared to peers who participated in basic early numeracy activities such as identifying numerals and verbal counting (Siegler & Ramani, 2009;

Whyte & Bull, 2008). Ramani and Siegler (2008) maintenance findings showed positive effects when using number line board games to develop early numeracy knowledge and skills. Student improvements in early numeracy knowledge and skills remained present two months after the conclusion of the intervention, and additionally, further benefits were shown on the future, more advanced arithmetic tasks (Siegler & Ramani, 2009). The evidence outlined above supports the use of number line board games as an effective support for early numeracy intervention for students with or at-risk for mathematics difficulties.

To implement number line board games requires only minimal training time, costs (i.e., materials), and prerequisite mathematic skills (e.g., Ramani, Siegler, & Hitti, 2012). Considering these basic requirements, the ease of implementation may allow non-teachers, such as older students (i.e., cross-age tutors) to deliver the intervention effectively, with high levels of fidelity, allowing for teachers to arrange smaller instructional groupings of students in need of individualized, explicit feedback and more frequent opportunities to practice mathematics skills (Clarke et al., 2017; Doabler et al., 2017).

Students with EBD and Peer Mediated Strategies

The academic and behavioral needs of students with emotional-behavioral disorders (EBD) have been identified as some of the most challenging to address (Kern, 2015). Frequently observed problem behaviors of students with EBD include struggles with peer acceptance (Ferguson, 1999), aggression, defiance (Gresham, Lane, MacMillan, & Bocian, 1999; Kauffman, 2001; Walker et al., 1995), off-task and other challenging behavior that can negatively influences both either own social-emotional and academic development, as well as that of their peers (Gunter, Denny, Jack, Shores, &

Nelson, 1993; Gunter et al., 1994). For students with EBD, each of these challenging behaviors can manifest themselves into detrimental short- and long-term outcomes.

Emotional-behavioral stressors experienced by students with EBD have been related to physical symptoms (e.g., headaches), anxiety, low self-esteem, disruptive classroom behavior, peer/teacher rejection, and low academic achievement (Reijntjes, Kamphuis, Prinzie, & Telch, 2010; Reijntjes et al., 2011; Soulis & Floridis, 2010). Long-term, students with EBD have the highest rate of dropping out of school than any other category of disability (Wood & Cronin, 1999). The limited number of students with EBD who do graduate from high school, rarely attend any form of postsecondary education (Malmgren, Edgar, & Neel, 1998), struggle with interpersonal relationships and are challenged when adjusting to vocational expectations (Gresham et al., 1999; Ollendick, Weist, Borden, & Greene, 1992; Walker et al., 1995). These detrimental behaviors show themselves in employment rates reported as low as 25% to just above 50% (Frank & Sitlington, 1997; Wagner, 1995). Seeking to address the range of needs of students with EBD, research has suggested that the dynamic between instruction and problem behavior can be utilized to deter disruptive classroom behavior, promote prosocial strategies, and increase the likelihood of positive academic outcomes (Deno, 1998; Gunter & Coutinho, 1997; Gunter & Denny, 1998; Wehby et al., 1998; Yoon, Barton, & Taiariol, 2004).

The issues and challenges related to the academic and behavioral needs of students with EBD have to be met through effective academic planning and selection of instructional techniques that result in positive outcomes for each student (Hughes & Fredrick, 2006; Vaughn et al., 2001). Additionally, the high priority placed on special educators' planning and instructional time has demanded a need for instructional techniques that are practical, low- or no-cost, and above all, effective (Bettini, Kimerling, Park, & Murphy, 2015; Blanton, Sindelar, & Correa, 2006; Brownell, Ross, Colon, &

McCallum, 2005; Greenwood, Carta, & Hall, 1988). Furthermore, when providing intensive intervention to students with disabilities, reducing the size of the instructional group or providing one-on-one tutoring can increase the intensity of instruction (Scammacca et al., 2007). With these considerations in mind, special educators may need to focus on underutilized resources within their own schools in order to provide individualized instruction that meets both the academic and non-academic needs of their students.

Peer mediated strategies

Peer-mediated strategies are frequently overlooked as effective, evidence-based instructional supports that can assist both new and experienced educators (Heron, Welsch, & Goddard, 2003). Peer mediated instruction has been used school settings to promote skills in reading, language arts, mathematics, social studies, and science (Calhoon, 2005; Hughes & Fredrick, 2006; Mastropieri et al., 2006; Mastropieri, Scruggs, Spencer, & Fontana, 2003). Furthermore, when implemented with rigor, peer-mediated intervention can produce positive collateral outcomes such as the maintenance and generalization of skills across settings, and when compared to adult-mediated intervention, may produce larger effect sizes (Jun, Ramirez, & Cumming, 2010; Kohler, Strain, Hoyson, & Jamieson, 1997; Strain & Kohler, 1999). Over several decades of research, peer tutoring has developed in terms of the models and the types of students serving in the roles of tutor and tutee. Variations on the model that have students with disabilities as participants include class-wide peer tutoring (Greenwood, Delquadri, & Hall, 1989), reciprocal (Fuchs, Fuchs, & Kazdan, 1999; Hughes & Frederick, 2006), reverse-role (Utley & Mortweet, 1997), and the focus of this study, cross-age tutoring (Heron, Welsch, & Goddard, 2003; Jun, Ramirez, & Cumming, 2010).

Cross-age tutoring

Cross-age tutoring utilizes an older student as the more knowledgeable and experienced peer in a coaching or instructional role, called the 'tutor', while the students receiving coaching or instruction from a tutor are called 'tutees' (Topping, 1998). An increasing research base in cross-age tutoring suggests that it can be an effective model for teaching academic and social skills to students with disabilities (Okilwa & Shelby, 2010; Robinson, Schofield, & Steers-Wentzell, 2005; Spencer 2006; Spencer, Simpson, & Oatis, 2009). Heron, Welsch, and Goddard (2003) identify cross-age tutoring as an intervention that requires minimal costs (i.e., time and materials) and can be implemented without substantial training time. Given the academic and behavioral challenges of students with EBD, utilizing the cross-age tutoring could be a way to provide direct, individualized instruction as well as provide opportunities for the tutor with EBD to practice and develop social, behavioral, and academic skills in an instructional setting.

Johnson and Bailey (1974) provided an early example of the cross-age tutoring model's utility in promoting kindergarten students' early numeracy knowledge and skills. Fifth-grade students were trained on, and role-played, instructional techniques, data collection procedures, and behavioral reinforcement strategies for use during individual tutoring sessions. After implementation of tutoring sessions, the tutees' improvements from pre- to post-test were significantly better than that of the control students who received no cross-age tutoring. Seeing as systematic and explicit instruction, and small or individualized instructional groupings are frequently found as cornerstone components of effective interventions for students with mathematics difficulties as well as the cross-age tutoring model, these mechanisms appear to be available to manipulation in order to increase treatment intensity (Bryant et al., 2011; Clarke et al., 2017; Gersten et al., 2009; Kroesbergen & Van Luit, 2003).

Blake, Wang, Cartledge, and Gardner (2000) found that with proper training, adult supervision and appropriate support, students with challenging behaviors and EBD can be effective instructors of social skills for younger students, and may also obtain collateral positive, non-academic benefits through this role. This role may be appropriate for students with EBD due to the unique interactions and behavioral requirements that role provides. Compared to teacher-led instruction, similarly aged students share a common social background with one another and can interact without the direct association of authority or control (Gaustad, 1993; Topping, 1996; Topping & Ehly, 1998), and additionally, the cross-age model alleviates the concern that students, especially those with emotional-behavioral challenges, may be intimidated or unwilling to learn from same-age peers due to risk of embarrassment (Gaustad, 1993). The cross-age model with students with EBD serving as tutors shows promise as an effective instructional model for promoting academic and/or behavioral skills for both the tutee and the tutor (e.g., Blake et al., 2000; Cochran, Feng, Cartledge, & Hamilton, 1993; Gumpel & Frank, 1999; Lane, Pollack & Sher, 1972; Lazerson, 1980; 2005; Maher, 1982; 1984; Scruggs & Osguthorpe, 1986; Top & Osguthorpe, 1987).

PURPOSE AND RESEARCH QUESTIONS

To date, there have been limited studies utilizing the cross-age tutoring model with students with EBD as tutors to address mathematics needs in the early grades (e.g., Robinson, Schofield, & Steers-Wentzell, 2005; Watts, Bryant, & Carroll, 2017). Furthermore, evidence has shown the effectiveness of number line board games in promoting early numeracy knowledge and skills in children at-risk for mathematics difficulties but research has not determined the feasibility, social validity, and overall effectiveness of this intervention when delivered by non-researchers or teachers, such as cross-age tutors (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008).

The purpose of this dissertation research was to investigate the effects of a cross-age tutoring model with tutors with EBD delivering instruction through number line board games to children (i.e., tutees) with mathematics difficulties. The study utilized a multiple baseline design across two participant populations (i.e., tutees and tutors), to determine the effects of the instructional model upon early numeracy performance of the tutees and classroom behavior of the tutors across their daily instructional classes (i.e., outside of the tutoring sessions).

The following research questions guided this study:

(1) What effect will a number line board game have on the mathematical performance of kindergarten students at-risk for mathematics difficulties, when delivered by older tutors (i.e., 5th-6th graders) with emotional-behavioral disorders?

(2) To what extent can students with emotional-behavioral disorders effectively serve as cross-age tutors and deliver early numeracy instruction through number line board games, as measured by the fidelity of implementation of tutoring procedures and tutees' outcomes on early numeracy measures?

(3) What effect(s) will the tutor training and implementation of the cross-age model have on the tutors' (with EBD) Check-in/Check-out behavioral scores across their instructional class periods (i.e., outside of the tutoring sessions)?

(4) To what extent do tutees maintain their mathematics performance and do tutors maintain their Check-in/Check-out behavioral scores, two and four weeks after the last tutoring session?

(5): To what extent are mathematical performance and classroom behaviors generalized, as measured by a distal measure (Test of Early Mathematics Ability-3;

Ginsburg & Baroody, 2003; and Student Risk Screening Scale; Drummond, 1994), for tutees and tutors, respectively?

(6) What are the perspectives of tutees, tutors, and participating special education teachers towards the cross-age tutoring intervention program?

Chapter 2: Literature Review

Rationale

Intervention studies targeting children at-risk for mathematics difficulties have focused on varying skills within the domain of early numeracy. Examples of early numeracy knowledge and skills addressed through mathematics intervention can be seen in the areas of vocabulary or mathematics language (e.g., Kleemans, Segers, & Verhoeven, 2011; Schleppegrell, 2010); classification, comparison, and reasoning (e.g., Pasnak et al., 2009); mathematical structures and symbol recognition (e.g., Andres, Di Luca, & Pesenti, 2008; Zhou & Wang, 2004); calculations (e.g., Carruthers & Worthington, 2004); measurement and geometry (e.g., Clements & Sarama, 2011); number line knowledge (e.g., Siegler, 2009); and counting knowledge (e.g., Askew, Bibby, & Brown, 2001; Gelman & Gallistel, 1978; Fuchs et al., 2010). This dissertation study utilized number line board games as the instructional materials and procedures for supporting the development of early numeracy knowledge and skills in children at risk for mathematics disabilities. Number line board games have been used in previous interventions for this population of students and have been shown to be effective in promoting skills in areas of counting, number comparison, number magnitude, and number line estimation (Aunio, Hautamäki, & Van Luit, 2005; Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008).

Number Line Board Games as Early Numeracy Intervention

Siegler and Booth (2004) proposed that playing a number line board game such as Chutes and Ladders supports knowledge of numerical magnitudes based on the rationale that this type of game shares visual, kinesthetic, auditory, and temporal connections to the number line system. The layout of these number line board games shows numbers arranged in equal-sized squares from left to right to support connections between the numbers and the magnitudes they represent. After spinning a number to identify how many spaces they may move forward, children are required to count-on from current number (i.e., space/location) on the board instead of the typical counting-from-1 procedure.

Siegler and Booth (2005) proposed a theoretical background for number line board games suggesting number line estimation as a developmental skill for early numeracy knowledge, due to the requirement of assessing magnitudes and utilize other early numeracy skills such as counting, number line estimation, number identification, and arithmetic (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008). Research has shown that early numeracy knowledge and skills are important to support the future development of mathematical skills (Jordan et al., 2009)

Recent research has focused on number line board games' effects on the early numeracy knowledge of children at-risk for mathematics disabilities. Children in preschool and kindergarten, from at-risk, low-socioeconomic backgrounds, are frequently the target populations of these studies. The number line board game materials varied only slightly depending on the age of the participating students. For example, for preschool students, the games contain numerals ranging from 0 – 10 (Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008), while for kindergarten participants, the numerals range from 0 – 100 (e.g., Laski & Siegler, 2014).

Intervention Studies for Children at-risk for Mathematics Difficulties

Siegler and Ramani have conducted multiple studies on the effects of number line board games on children at-risk for mathematics difficulties. Ramani and Siegler (2008) initially tested their theory and found that when preschool students from low-socioeconomic backgrounds played a number line board games for 15-20 m per day, 4 days per week, over 2 weeks, they significantly improved early numeracy knowledge and skills in the areas of counting, numerical magnitude comparison, number line estimation, and number identification, compared to peers who played a non-number, 'colored spaces' board game. Maintenance of these skills was measured after 9 weeks and effects were found to still be present for the students in the treatment group.

Siegler and Ramani (2008) found similar results when they evaluated the effects of playing number line board games versus colored spaces board games on preschool children's numeral magnitude knowledge. In experiment one, Siegler and Ramani (2008) identified a gap between the number magnitude skills of students from low-socioeconomic backgrounds and their peers from middle-income or more advantaged backgrounds. In experiment two, the preschoolers from low-income backgrounds attended four, 15 m sessions over two weeks, where they played the number line board games with a researcher. After the intervention, the gaps in number magnitude knowledge between students from low-socioeconomic backgrounds and their more advantaged peers were seen to disappear. Children who played the colored board games showed no improvement in numerical magnitude knowledge.

Next, Siegler and Ramani (2009) attempted to parse out the components within the number line board games that produced the improvements in early numeracy knowledge and skills in the earlier studies. The researchers hypothesized that the improvements in numerical competency would be greater when students played linear number line board games than when they played circular number board games due to the direct connection between the linear number line representation and the desired mental representation. Compared to preschoolers who played the circular board game or participated in other numerical activities (e.g., basic counting), the students who played the linear number board game showed greater increases in numerical magnitude comparison and number line estimation skills. Additionally, the children who played the number line board games performed better on follow-up training on arithmetic problem-solving tasks. These earlier findings were supported by Ramani and Siegler (2010), where preschool students who had less initial numerical competency were able to make greater gains on measures of number line estimation, magnitude comparison, numeral identification, and arithmetic learning, compared to peers with more advanced, initial knowledge, after playing number line board games.

Whyte and Bull (2008) also attempted to identify the differential effects of three intervention board games based on their design and numerical components. The three board games evaluated were a linear numerical board game, a linear (non-numerical) colored board game, and a non-linear (i.e., cards numbered 1-100) numerical game. Preschool aged children in each treatment condition were provided four 25 m sessions where they played the assigned board game. Both numerical game conditions significantly outperformed the non-numerical, colored board game condition on number estimation and counting abilities. Furthermore, preschoolers in the linear number board game significantly outperformed the non-linear (card) number game on tasks related to number identification.

Summary

Across the intervention studies for young children at-risk for mathematics difficulties, students who participated in number line board games showed improvements in numerical competency in the areas of counting, numerical magnitudes, number comparison, and number line estimation, compared to peers who engaged in basic early numeracy activities such as identifying numerals and verbal counting (Siegler & Ramani, 2009; Whyte & Bull, 2008). Maintenance of the positive effects of number line board games was shown by Ramani and Siegler (2008), where student improvements in mathematical knowledge two months after the conclusion of the intervention, and additionally, the further benefits of these games was shown on the development of more advanced mathematical learning in arithmetic skills (Siegler & Ramani, 2009). The evidence outlined above supports the use of number line board games as an effective early numeracy intervention for closing the numerical competency gap between students at-risk and their typically developing peers.

STUDENTS WITH EBD AS CROSS-AGE TUTORS

Role Theory and Cross-age Tutoring

The foundational theory for cross-age tutoring is based upon Piaget's research in socio-cognitive conflicts and their ability to promote learning (1977). The conflict develops when a learner is met with information that challenges his or her own perceptions or assumptions. When peer interactions and cooperative learning take place, children are able to obtain social and cognitive benefits such as the promotion of communication skills and the opportunity to identify and understand different perspectives (Damon & Phelps, 1989; Piaget; 1977).

Because student tutors are not professional educators, socio-cognitive conflicts can be also present during student tutorial learning as well (Roeders, 1995). Vygotsky's (1978) developmental theory is also applicable because it emphasizes the effect social-communicative interactions on learning. This point of view also attempts to explain how and why the tutor may be able to receive positive skill development from being in the role of tutor. Due to the social-communicative interactions within the model, the tutor is required to use deep cognitive processes in order to motivate, attend, and explain to the tutee (Vygotsky, 1978). Additionally, through interaction opportunities provided within the tutoring model, the tutor is exposed to responsibility, increased status, respect, and receives direct attention from the tutee(s) (Hogan & Tudge, 1999). This can promote interest in learning, and increase effort and intrinsic motivation (Allen, 1976). The theory suggested as a way of interpreting the mechanism producing these outcomes is known as 'role theory' (Allen, 1976; Bierman & Furman, 1981).

Role theory proposes that the behavior of the individual is influenced by the role they inhabit or play (Thomas & Biddle, 1966). Roles are defined through society's association of a certain set of attitudes or behaviors with a given identity, and that when an individual assumes that role, they begin to develop or align their own self-perceptions with that of the given role (Turner, 2002). When a student accepts the role of tutor, it is proposed that they are undertaking a role similar to that of a teacher, which requires a set of familiar behaviors (Foot, Shute, Morgan, & Barron, 1990). The role of a teacher/tutor requires attitudes and behaviors (e.g., teaching skills, active listening, answering questions, reinforcing behavior, corrective feedback) that function on more responsibility and independence than that which students have in the traditional teacher-student instructional model (Allen, 1976; Allen & Feldman; 1973, 1976). Robinson, Schofield, and Steers-Wentzell (2005) suggest that when students with or at-risk for disabilities undertake the role of a cross-age tutor, the large discrepancy between that role and their typical student identity requires a transformation that produces "spillover" effects seen in the form of increases in academic domains, time on-task, classroom behavior, and positive attitudes towards school (Allen & Feldman, 1976; Bierman & Furman, 1981). There are a few explanations that attempt to theorize the interactions within the cross-age tutoring model that may produce these outcomes.

First, within the tutoring sessions, the tutor is exposed to student behaviors that are the target of development or influence learning development. Being in the role of tutor and observing the learning process, as well as the challenges of teaching another student, the tutor may develop and understand and insight into the importance of student behavior (e.g., staying on task, listening, following directions) on learning. Essentially, they may be observing and identifying the essential skills or barriers to being a "good' student. Henceforth, when they return to the role of student, they attempt to utilize attitudes and behaviors that may be desired for their tutees (Robinson, Schofield, & Steers-Wentzell, 2005). The second explanation suggests that the cross-age tutor is developing or exhibiting skills that allow him or her to be seen as a good role model for their tutee due to their older age (Schunk, 1998). In order to be seen as a role model to their tutee and benefit from receiving the desirable positive attention of a younger student, the tutor may begin to develop positive social and behavioral skills to take on that positive role model identity with more frequency (Smead, 1984).

But why is the cross-age model appropriate for students with EBD and sometimes found to be even more effective than typical teacher-lead or same-age peer instructional models (Bowman-Perrott, Burke, Zhang, & Zaini, 2014; Jun, Ramirez, & Cumming, 2010)? Although educators have background training on instructional strategies and techniques that older students frequently do not, similarly aged students share a common social background with one another and can interact without the direct association of authority or control (Gaustad, 1993; Topping, 1996; Topping & Ehly, 1998). Alternatively, students may be intimidated or unwilling to learn from same-age peers due to risk of embarrassment (Gaustad, 1993), and therefore, cross-age tutoring may provide an instructional model, based upon these underlying theories, that may promote engagement and learning in the tutee, and positive academic and behavioral improvements for the tutor.

Blake, Wang, Cartledge, and Gardner (2000) noted that with proper training, adult supervision and appropriate support, students with challenging behaviors and EBD can be effective instructors of social skills for younger students, and may also obtain collateral positive, non-academic benefits through this role. Research focusing on students with challenging behaviors in the role of cross-age tutor has been limited in recent years, but has shown positive outcomes for the tutor, as well as for the tutee (Blake et al., 2000; Gumpel & Frank, 1999; Lazerson, 1980; Maher, 1982; 1984; Scruggs & Osguthorpe, 1986; Top & Osguthorpe, 1987; Watts & Bryant, 2017). Improvements in the areas of reading (Cochran, Feng, Cartledge, & Hamilton, 1993; Lane, Pollack & Sher, 1972; Top & Osguthorpe, 1987), mathematics (Franca, Kerr, Reitz, & Lambert, 1990; Robinson, Schofield, & Steers-Wentzell, 2005), spelling (Stowitschek, Hecimovic, Stowitschek & Shores, 1982), test scores, and grades (Maher, 1982; 1984) have been found for tutors with EBD. In addition to academic achievement, research on same-age and cross-age tutoring models also suggest positive outcomes in social, emotional, and behavioral skills, including discipline within the classroom setting and the reinforcement of peer relationships (Greenwood, Carta, & Hall, 1988; Maher 1982; 1984), social skills (Blake et al., 2000; Gumpel & Frank, 1999; Watts & Bryant, 2017), on-task behavior (Greenwood, Delquadri, & Hall, 1989; Hogan & Prater, 1993), self-esteem and self-worth (Lazerson, 2005; Miller, Topping, & Thurston, 2010), and attendance rates (Maher, 1982). Given that these social-behavioral, and previously mentioned academic skills are frequently characterized as deficit areas for individuals with EBD (Landrum, Tankersely, & Kauffman, 2003; Trout, Nordness, Pierce, & Epstein, 2003), utilizing cross-age peer tutoring as a possible intervention show promise to address these needs.

Relevant Reviews of Peer Tutoring with Students with EBD

Two early syntheses of the literature found that students with a range of disabilities can act as effective peer tutors across content and skill areas (Osguthorpe & Scruggs, 1986) and furthermore, being in the role of tutor may assist students with disabilities to develop positive attitudes and increase self-concept (Cook, Scruggs, Mastropieri, & Casto, 1985). A number of systematic reviews completed within the last few years have focused on both academic outcomes, and less frequently, social-emotional and behavioral outcomes, in regards to students with disabilities and peer-mediated interventions (Bowman-Perrott, Burke, Zhang, & Zaini, 2014; Bowman-Perrott, Davis, Vannest, Williams, Greenwood, & Parker, 2013; Okilwa et al, 2010; Ryan, Reid, & Epstein, 2004; Spencer, 2006; Spencer et al., 2009).

Most recently, Bowman-Perrott, Burke, Zhang, and Zaini (2014) conducted a meta-analysis of 20 studies focusing on direct and collateral effects of peer tutoring on social and behavioral outcomes for students with disabilities. Findings showed that peer tutoring had a greater effect on promoting social skills and reducing disruptive behaviors than increasing academic engagement for students with disabilities, and also, that cross-age tutoring was more effective than same-age or reciprocal tutoring for students with EBD. Similar findings were obtained by Jun, Ramirez, & Cumming's (2010) meta-analysis on tutoring in literacy, where cross-age tutoring was found to be more effective than adult tutoring and computer-based tutoring, especially when students with disabilities served as tutors.

Bowman-Perrott, Davis, Vannest, Williams, Greenwood, and Parker (2013) also examined peer tutoring effects on academic skills in a meta-analysis that included 26 single-subject design studies. Findings of this review showed the model to be highly effective for students in grades first through twelfth and that students with EBD obtained greater benefit from the model than other disability types. Ryan, Reid, and Epstein (2004) also focused their review on the academic achievement of peer-mediated interventions for students with EBD. Overall findings of the synthesis suggest that peer-mediated interventions appear to be effective across content areas for students with EBD. Additionally, findings suggest that research in the area of academic interventions for students with EBD has declined in recent years.

One of the most comprehensive reviews in this area was undertaken by Spencer (2006), which reviewed 37 studies from 1972 to 2002 with a focus on students with EBD within cross-age and same-age peer tutoring models. The author calculated quantitative outcomes for reading, spelling, mathematics and social studies, across the different peer tutoring arrangements. The cross-age tutoring model was noted as more effective than both the same-age and reciprocal tutoring (i.e., students alternate between tutor and tutee; Fantuzzo, Davis, & Ginsburg, 1995; Fantuzzo, King, & Heller, 1992) in reading, but less effective than the same-age tutoring model for mathematics. It should be noted that only 13 studies provided sufficient data to calculate effect sizes. Spencer and colleagues (2009) continued the previous review by identifying nine additional studies from 2001-2007 that included students with EBD in tutor and tutee roles within peer tutoring models, across elementary and secondary settings, however, mainly within special education classrooms. The authors noted that although peer tutoring continues to show promise as an effective intervention for students with EBD as tutors or tutees, additional research is required for these students in secondary and generalized settings.

Intervention Studies with Students with EBD as Cross-age Tutors

Eleven intervention studies utilizing tutors with EBD within the cross-age model have been conducted between the years of 1972 through 2016. The earliest study was conducted by Lane, Pollack, and Sher (1972), and contained tutors with EBD in grades eight and nine, and tutees with learning disabilities and behavioral disorders in grades three and four. Tutors were trained in social skills and literacy instructional methods and delivered instruction for reading, writing, and spelling skills twice a week for a duration of 7 months. Both tutees and tutors increased their reading achievement scores at statistically significant levels, and the tutors showed large effects (i.e., significant decreases) on a measurement of disruptive behaviors.

Lazerson (1980) trained students with aggressive and withdrawn behaviors in grades five through eight to tutor students with similar behavioral characteristics in grades two through five. Tutors were trained in modeling, role-playing, positive reinforcement and corrective feedback techniques, but were given the ability to choose, structure, and manipulate the content of each tutoring session at their own discretion. Tutoring took place for 20-30 m per day, five times per week over ten weeks. Tutors and tutees showed statistically significant gains on behavioral and self-concept measures. Of the studies that measured the social validity of utilizing cross-age tutoring with tutors with EBD, this was one of the rare studies where the teachers had mixed perceptions of the benefits for the participating students due to the unstructured nature of the tutoring sessions. Although, the teachers did report a high level of interest in continuing the tutoring program if changes to the structure were made.

Maher (1982; 1984) conducted two similar studies utilizing high school cross-age tutors with EBD to provide instruction to elementary and middle school students with learning or intellectual disabilities. In both studies, a special educator and school psychologist/counselor implemented the tutoring program. Both studies provided training to tutors on lesson planning, instructional techniques, evaluation methods, and problem-solving. In each study, tutoring sessions took place for 30 min, two days per week, ten weeks. Instructional content included reading, writing, and mathematics skills. In the first study (Maher, 1982), tutors' outcomes were statistically significant compared to non-tutor peers in the areas of disciplinary referrals, grades, and percentage of school days in attendance. Tutee outcomes were not measured. In the second study (Maher, 1984), large effect sizes were shown on tutees' and tutors' percentage of assignments completed, test and quiz grades, and disciplinary referrals. Additionally, the tutors' fidelity of implementation of tutoring procedures was also evaluated and found to be consistently high across tutors (M = 96.3%).

Scruggs and Osguthorpe (1986) provided tutor training to students with EBD in grades two through six to provide instruction to younger students in grades one through five, utilizing a structured reading curriculum. Tutoring sessions took place for 30 m, two to five days per week, over ten weeks. Tutees and tutors did not show statistically significant improvement in performance on the Woodcock-Johnson compared to the control group. On another reading assessment, aligned with the reading curriculum used during the tutoring sessions, the tutors and tutees showed statistically significant improvement in performance compared to the control group. On an attitude towards school measure, tutees showed statistically significant improvement compared to control students, while tutors showed little change compared to non-tutor peers.

Top and Osguthorpe (1987) trained a teaching assistant to implement a cross-age tutoring program with tutors with EBD and learning disabilities in grades fourth through sixth and tutees at-risk for reading disabilities in first grade. Tutors were trained on modeling, prompting, positive reinforcement and progress monitoring techniques for delivering instruction and assessment on reading skills (e.g., phonics, sight words). Tutoring sessions took place for 15-20 m, four days per week, over 14 weeks. Tutees showed mixed outcomes on reading measures, from no effects to large, statistically significant improvements compared to the control group. Tutors showed statistically significant improvement compared to the control group on the reading measure, but small to no effect on various self-concept and perception of ability measures.

Cochran, Feng, Cartledge, and Hamilton (1993) trained fifth-grade students with or at-risk for EBD and struggling readers, to tutor second-grade students with or at-risk for EBD and struggling readers. Tutoring sessions took place for 28-30 m per day, for 32 total sessions, over 8 weeks. During tutoring sessions, the tutors provided instruction, practice opportunities, review games, reinforcement, and evaluated and charted student progress. On the sight words curriculum-based measure, tutors and tutees showed small improvements compared to students in the control groups. On a standardized, behavioral measure completed by teachers, tutors showed moderate to large improvements across measures of social skills, problem behaviors, and academic skills. On the same measure, tutees showed no differences compared to the control group. On a self-assessment version of the same measure, tutees showed moderate improvements and tutors showed no effects. Tutor dyads were also directly observed to assess the frequency of positive (e.g., cooperative statements) and negative behaviors (e.g., 'put-down' statements) during tutoring sessions. On average, tutoring dyads showed overall increases in positive behaviors and decreases in negative behaviors during tutoring sessions.

Hogan and Prater (1993) studied the effects of a 15-year-old cross-age tutor with EBD on a 14-year-old tutee with a learning disability. Tutoring instruction was provided in spelling and vocabulary skills for 15 m sessions, four days per week. The tutee showed large effects from baseline to intervention phase on spelling test scores, vocabulary test scores, and on-task behavior. The tutor was assessed for changes in the frequency of disruptive behaviors and showed small effects in decreasing target behaviors from baseline to intervention phase.

Grumpel and Frank (1999) studied the sixth-grade cross-age tutors who were at-risk for EBD and their effect on increasing positive interactions of kindergarten students at-risk for EBD. The tutors were trained on an instructional procedure that components that support modeling, role playing, and the use of a self-monitoring sheet. Tutoring sessions took place four times per week, and researchers measured the frequency of positive social interactions and 'no social interactions' through momentary time sampling. The tutees and tutors both showed large effects in increasing positive social interaction and decreasing no social interactions during the sessions. Maintenance of these large effects was present at the same levels for both tutors and tutees on follow-up measures.

Blake, Wang, Cartledge, and Gardner (2000) trained middle school students with EBD to implement cross-age tutoring with a scripted social skills curriculum to third and fourth-grade students who were identified as having difficulties relating to peers. Tutoring sessions took place for 45 m each day, five days per week, over seven weeks. During the tutoring sessions, tutors would model, practice, play informal games, and provide feedback with their tutee. Both the tutees and the tutors were directly observed to assess the frequency of supportive and abusive behaviors. Tutees and tutors both showed large effect sizes in increasing supportive behaviors and decreasing abusive behaviors. When maintenance of these effects was evaluated after the conclusion of the tutoring program, the effect sizes for both tutees' and tutors' target behaviors remained large. Additionally, tutors implemented instructional procedures with consistently high rates of fidelity (M = 97%).

Lazerson (2005) trained 15- and 16-year-old students with EBD to tutor first- through fifth-grade students with Individualized Education Programs. Tutor training consisted of practice and role-playing activities, and corrective reinforcement statements.

Tutoring sessions took place two to five times per week for 45-60 m, over a duration of three months. Tutors provided instruction on reading comprehension, decoding, and mathematics skills. Tutors were assessed for changes in self-concept. Tutors showed moderate effects in improving self-concept from pre- to posttest. Tutee outcomes were not assessed in this study.

Summary

The literature base for tutors with EBD serving within the cross-age model provides a number of key insights. First, findings show that students with EBD are capable of serving as cross-age tutors, as they show consistently high rates of implementing tutoring procedures with fidelity. Second, the cross-age tutoring model with tutors with EBD shows consistently moderate to large effects, on average, for tutees and tutors in academic outcomes and behavioral skills. Third, practitioners (e.g., teachers, paraprofessionals, school psychologist) were rarely the primary implementers of the tutoring program (e.g., Maher, 1982; 1984; Tops & Osguthorpe, 1987), which shows a need for research to be undertaken to assess the feasibility, social validity, and overall effectiveness in clinical applications. This is an important next step for this line of research as teachers who implement and perceive a practice as beneficial, they are more likely to continue implementing that practice or program (Zimmerman & Risemberg, 1997). Fourth, the maintenance and generalization of the skills targeted for tutoring instruction and related outcomes, usually for the tutor with EBD, are measured inconsistently across the empirical literature. Additionally, there have been few examples of the effectiveness of utilizing the cross-age tutoring model with tutors with EBD to promote early numeracy performance (e.g., Holecek, 2012), and there has yet to be a study that has used the model in combination with number line board game instructional procedures and materials, a hole in the literature base that this dissertation study fills.

Chapter 3: Method

The purpose of this study was to identify the effectiveness and related outcomes of utilizing cross-age tutors (i.e., older students in grades five and six) with emotional-behavioral disorders (EBD) to promote early numeracy performance in young children (i.e., ages 3-6) at-risk for mathematics difficulties, through playing number line board games 3 days per week, for 25-30 m sessions, over 10 weeks. The research questions for the study were: (1) What effect will a number line board game have on the mathematical performance of kindergarten students at-risk for mathematics difficulties, when delivered by older tutors (i.e., 5th-6th graders) with emotional-behavioral disorders? (2) To what extent can students with emotional-behavioral disorders effectively serve as cross-age tutors and deliver early numeracy instruction through number line board games, as measured by the fidelity of implementation of tutoring procedures and tutees' outcomes on early numeracy measures? (3) What effect(s) will the tutor training and implementation of the cross-age model have on the tutors' (with EBD) Check-in/Check-out behavioral scores across their instructional class periods (i.e., outside of the tutoring sessions)? (4) To what extent do tutees maintain their mathematics performance and do tutors maintain their Check-in/Check-out behavioral scores, 2 and 4 weeks after the last tutoring session? (5): To what extent are mathematical performance and classroom behaviors generalized, as measured by a distal measure (Test of Early Mathematics Ability–3; Ginsburg & Baroody, 2003; and Student Risk Screening Scale; Drummond, 1994), for tutees and tutors, respectively? (6) What are the perspectives of tutees, tutors, and participating teachers towards the cross-age tutoring intervention program? See Table 3.1 for research question alignment to dependent variables and measures.

Research Question	Dependent Variables	Measure
Research question 1: Tutee Outcomes What effect will a number line board game have on the mathematical performance of kindergarten students at-risk for mathematics difficulties, when delivered by older tutors (i.e., 5th-6th graders) with emotional-behavioral disorders?	• Tutees mathematics performance on weekly proximal measure of early numeracy skills	• TEMI–AC; administered weekly
Research question 2: Tutors with EBD To what extent can students with emotional-behavioral disorders effectively serve as cross-age tutors and deliver early numeracy instruction through number line board games, as measured by the fidelity of implementation of tutoring procedures and tutees' outcomes on early numeracy measures?	• Outcomes related to research question 1 (i.e., effectiveness in promoting mathematics performance of tutees) • Fidelity of implementation of tutoring procedures	• TEMI–AC scores for tutees • Fidelity of implementation of tutoring procedures checklist – Completed for 30% of tutoring sessions • Number of tutor re-training sessions required for each tutor
Research question 3: Tutor Outcomes What effect(s) will the tutor training and implementation of the cross-age model have on the tutors' (with EBD) ability to obtain Check-in/Check-out behavioral points across their instructional class periods (i.e., outside of the tutoring sessions)?	• Tutors: Classroom behavior scores for instructional periods	• CICO completed by teachers for each tutor after each instructional period, daily
Research question 4: Maintenance To what extent do tutees maintain their mathematics performance and do tutors maintain their Check-in/Check-out behavioral scores, 2 and 4 weeks after the last tutoring session?	• Tutees: Mathematics performance on early numeracy proximal measure • Tutors: Classroom behavior scores for instructional periods	• TEMI–AC (administered 2 and 4 weeks post-intervention) • CICO (continually scored for 2 and 4 weeks post-intervention)
Research question 5: Generalization To what extent are mathematical performance and classroom behaviors generalized, as measured by a distal measure, for tutees and tutors,	• Tutees: Mathematics performance on early numeracy distal measure	• Tutees: Re-administration of TEMA-3, post-intervention

respectively?	• Tutors: Risk level/category for EBD on social-behavioral rating scale	• Tutors: Re-administration of SRSS, post-intervention
Research question 6: Social Validity What are the perspectives of tutees, tutors, and participating teachers towards the cross-age tutoring intervention program?	• Teachers: Practicality, feasibility, and perceived benefits/effectiveness of program • Tutors: Perceived benefits/effectiveness, strengths and challenges of program • Tutees: Perceived benefits/effectiveness, strengths and challenges of program (*Questions dictated; responses recorded)	• Researcher-developed social validity survey containing rating scales and open-ended questions

Note. TEMI–AC = Texas Early Mathematics Inventory–Aim Checks (University of Texas System/Texas Education Agency, 2009); TEMA–3 = Test of Early Mathematics Ability–3 (Ginsburg & Baroody, 2003); CICO = Check-in/Check-out; SRSS = Student Risk Screening Scale (SRSS; Drummond, 1994)

Table 3.1: Research Questions, Dependent Variables, and Measures.

University and school IRB approval, as well as parent or guardian consent for potential tutees and tutors, was obtained prior to the start of the study. Assent forms were also collected for potential tutor participants. School administrators and collaborating teachers were provided with the aims of the study and contacts were established with both of the participating teachers (i.e., kindergarten teacher and cross-categorical teacher for grades 3–6). All teachers voluntarily agreed to complete all training, implementation, and assessment requirements of the study.

PARTICIPANTS

Screening procedures (see: Measures section) yielded five tutees at-risk for mathematics disabilities and five tutors with or at-risk for emotional-behavioral disorders (EBD). Tables 3.2 and 3.3 provide student demographic and screening test information for tutees and tutors, respectively. Note that tutor/tutee dyads are aligned with their given number (i.e., tutee 1 and tutor 1 formed a dyad, tutee 2 and tutor 2 formed a dyad).

Tutees

Five tutees qualified for the intervention. All were males attending general education kindergarten, ages 5 years to 5 years-5 months (See Table 3.2). Three of the students were Caucasian, one student was African-American, and one student was Hispanic. Four of the students qualified for free/reduced lunch due to low socioeconomic status. All students were identified by their teacher as having difficulties in the area of early numeracy knowledge and skills, and possibly in need of intervention. Screening results from the administration of the TEMA-3 showed one student's score to fall within the 'below average' range and the remaining four of the students' scores fell under the 'poor' category. None of the students were currently receiving additional support for mathematics difficulties prior to the intervention, although one student (i.e., Tutee 5) was being assessed for difficulties related to attention and behavioral concerns.

Note. K = kindergarten; M = male; IEP = Individualized Education Program; Y = yes; N = no; TEMA–3 = Test of Early Mathematics Ability–3 (Ginsburg & Baroody, 2003); * = Raw score/Math Ability score

		Tutee 1	Tutee 2	Tutee 3	Tutee 4	Tutee 5
Age (years-months)		5-3	5-2	5-5	5-3	5-0
Grade		K	K	K	K	K
Gender		M	M	M	M	M
Race/Ethnicity		African-American	Hispanic	Caucasian	Caucasian	Caucasian
IEP/ Disability status		None/ At-risk	None/ At-risk	None/ At-risk	None/ At-risk	None/ At-risk
Free/Reduced lunch		Y	Y	Y	Y	N
Screening: TEMA-3	**Raw / MA*	10 / 77	8 / 77	11 / 80	9 / 75	6 / 72
	Category	Poor	Poor	Below average	Poor	Poor

Table 3.2: Tutees' Demographic and Screening Test Information.

Tutors

Table 3.3 displays tutor demographic and screening information. The five students who qualified to be tutors due to their IEP disability category and/or risk status for EBD were in fifth- (i.e., three students) and sixth-grade (i.e., two students). Students' ages ranged from 10 years-9 months to 12 years-2 months. All of the tutors were male and four of the five students qualified for free/reduced lunch due to low socioeconomic status. Three of the students had IEPs with EBD as the designated, primary disability category, while the remaining students had learning disabilities and were also perceived to be at-risk for EBD due to challenging behaviors, as identified by their special education teacher and case manager. Screening scores on the SRSS indicated moderate- to high-risk for internalizing and externalizing disorders, across all students. Disaggregated raw scores and risk-status categorization can be found in Table 3.3.

		Tutor 1	Tutor 2	Tutor 3	Tutor 4	Tutor 5
Age (years-months)		10-9	11-6	10-5	12-2	11-0
Grade		5	6	5	6	5
Gender		M	M	M	M	M
Race/Ethnicity		Caucasian	Caucasian	Caucasian	Caucasian	Caucasian
IEP disability category		EBD	LD & at-risk for EBD	EBD	LD & at-risk for EBD	EBD
Free/Reduced lunch		Y	N	Y	Y	Y
Screening: SRSS score (EBD risk status)	*External*	10 (High)	11 (Moderate)	17 (High)	13 (High)	15 (High)
	Internal	4 (Moderate)	8 (Moderate)	13 (High)	13 (High)	11 (High)

Note. M = male; IEP = Individualized Education Program; EBD = Emotional-behavioral disorder; LD = learning disability; Y = yes; N = no; SRSS = Student Risk Screening Scale (SRSS; Drummond, 1994); External = externalizing disorder; Internal = internalizing disorder

Table 3.3: Tutors' Demographic and Screening Test Information.

Teachers

With the permission of the district and the school principal, school staff were contacted and recruited to implement the cross-age tutoring model. Brief recruitment meetings were conducted online (i.e., via Skype) where the researcher outlined the intervention in detail, including teacher responsibilities, assessment requirements, and targeted outcomes for each of the student populations. Preliminary questions were answered and a brief, written overview of the details of the intervention and timeline was provided to the teachers and administrators. After the initial meetings, both teachers agreed to participate in the training and implementation requirements of the study. These practitioner implementers were selected based on their familiarity with the participants (i.e., they were the primary provider of students' instructional minutes throughout the school day and/or were the students' IEP case managers). Both teachers had state licensed teaching certificates. The teacher of the cross-age tutors was a Caucasian male and taught third through sixth-grade special education in a self-contained, cross-categorical classroom, serving students with IEP labels of emotional-behavioral disorders and learning disabilities. This teacher had 9 years of lead teaching experience in elementary and middle school special education settings. The teacher of the tutees was a Caucasian female and taught kindergarten students, ages 5-6. She had 12 years of lead teaching experience in the general education kindergarten setting as well as a master's degree in the same area. Both teachers had been teaching at the site school for at least three years.

SETTING

The study took place at a public elementary school in a suburban school district in central Colorado. The school served 444 students in grades preschool through sixth-grade, and students attended on a year-round school schedule. School records from 2016 showed that 40.4% of the students qualified to receive free or reduced-price lunch. The special education population comprised 10.7%, English language learners represented 16.2%, and gifted students were identified at 3.3% of the school population. School demographics in ethnicity show 58.4% of students identifying as Caucasian, 25.4% identifying as Hispanic or Latino, 14.5% of students identifying as other ethnic minority, and less than 3% identifying as African-American.

All intervention tutoring sessions occurred within the elementary school. The classroom utilized for the tutor training and intervention tutoring sessions was that of the upper elementary special education teacher (grades three through six). The tutoring sessions took place 3 days per week for 25-min, over the duration of 10 weeks. The sessions took place at the same time each day (i.e., 10:30am) during the same days of the week (i.e., Tuesdays, Wednesdays, and Thursdays), with the exception of days when assemblies and fire drills were scheduled. The sessions that occurred on these days were rescheduled to an afternoon time or took place on a Friday morning at the regularly scheduled time.

EXPERIMENTAL DESIGN

A single-case, co-occurring multiple baseline design across participants (Kennedy, 2005) was implemented to evaluate the effects of a cross-age tutoring program with tutors with EBD on the mathematics performance of children with mathematics difficulties. The basis of single-case research methodology relies upon repeated measurement of dependent variables before, during, and after the introduction of the independent variable to determine if a causal relation exists (Horner et al., 2005; Kennedy, 2005). The single-case design is especially effective for populations with unique characteristics that less readily available/accessible for research participation (Horner et al., 2005; Kratochwill et al., 2010), such as the students with EBD included in this study. Furthermore, this design is desirable for students with intensive needs due to a methodology that does not require withdrawing the potentially beneficial intervention in order to determine effectiveness (Kennedy, 2005).

Multiple baseline designs typically include the establishment of concurrent baselines, the observation of stable baseline levels, and the sequential introduction of the independent variable (e.g., intervention) across participants (Horner et al., 2005; Kennedy, 2005; Kratochwill et al., 2010). Single-case designs require the establishment of experimental control through the determination of functional relation(s) (Kratochwill et al., 2010). A functional relation is defined as the consistent effect on the dependent variable through systematic manipulation of the independent variable (Kennedy, 2005). The functional relation can be observed as a change in the dependent variable after the independent variable is introduced, and the experimental control determined when the effect is consistently observed after the independent is systematically introduced across participants (Horner et al., 2005; Kratochwill et al., 2010). The more replications of the relation across participants increases the evidence for experimental control (Kratochwill & Levin, 2010).

For this study, two sets of concurrent multiple baseline designs were implemented, one set for the tutee participants, and one set for the tutor participants. The tutees' dependent variable was mathematical performance (i.e., total scores) on the Texas Early Mathematics Inventory–Aim Checks (TEMI–AC; University of Texas System/Texas Education Agency, 2009), and the tutors' dependent variable was measured through weekly averaged scores on Check-in/Check-out behavioral point sheets. The independent variable for tutees was attending cross-age tutoring sessions in which they participated in number line board games for 20-25-min per day, 3 days per week, over 10 weeks. The independent variable for the tutors included two components: (1) the tutor training sessions in which they received instruction on tutoring skills, number line board game procedures, and positive behavioral reinforcement strategies; and (2) implementing the cross-age tutoring program through individual tutoring sessions with their tutee.

After a stable baseline has been identified, the tutoring program (i.e., intervention) is systematically introduced to each tutor dyad. The effect on the dependent measures will determine the presence of a functional relation. The evidence for experimental control will be exhibited through the replication the effects on the dependent measures across participants (i.e., tutees and tutors).

Quality indicators

To ensure rigor, this dissertation study followed the quality indicators for single-subject experimental designs, as outlined by Horner and colleagues (2005). Quality indicators for single-subject designs fall within the following guideline categories: (a) experimental design, (b) participants and settings, (c) dependent variable, (d) description and manipulation of independent variable, (e) baseline, (f) experimental control, (g) external validity, and (h) social validity (Horner et al., 2005; Kratochwill et al., 2013). The purpose for utilizing these quality indictors is to deter potential threats to internal and external validity (Kazdin, 2011; Kennedy, 2005).

Independent Variables

The intervention contained two components that constituted the independent variables. The first component was the training of cross-age tutors with EBD on the following tutoring skills: (a) number line board game procedures and corrective feedback techniques (adapted with permission from Ramani, Siegler, & Hitti, 2012); (b) positive behavioral support statements and strategies; and (c) supervision techniques for keeping tutees on-task during the administration of weekly progress monitoring measures. The second component was the implementation of cross-age tutoring sessions in dyads, utilizing number line board games for 25-30-min sessions, 3 days per week, over 10 weeks.

Dependent Variables

Table 3.1 displays the dependent variables aligned with each of the research question, and the related measures.

Research question 1: Tutee outcomes

The dependent variable for the first research question is the tutees', with mathematics difficulties, mathematical performance. The individual tutee's performance was measured through the administration of a weekly progress monitoring assessments during the baseline, intervention and maintenance phases (i.e., TEMI–AC). Effectiveness of the cross-age tutoring program was evaluated through visual analysis of graphical data procedures (Horner et al., 2005; Kennedy, 2005; Kratochwill et al., 2010; Parsonson & Bear, 1978) and the calculation of effect sizes (i.e., percentage of data points exceeding the median of the baseline phase; PEM; Chen & Ma, 2007).

Research question 2: Effectiveness of cross-age tutors with EBD

To answer the second research question, related to assessing the extent to which students with EBD can effectively serve as cross-age tutors, the evaluation of two dependent variables is required. The first dependent variable assessed was the fidelity of tutors' instruction during tutoring sessions. This dependent variable was the defined as the extent to which the cross-age tutors with EBD follow and implement tutoring procedures during tutoring sessions. The tutoring procedures are directly aligned with the components and skills embedded and evaluated within the tutor training sessions provided by the investigator, immediately prior to the start of tutoring sessions (i.e., intervention phase). This dependent variable was measured through a fidelity of implementation of tutoring procedures checklist during 30% of the tutoring sessions, for each cross-age tutor. The results of the fidelity measure in combination with the outcomes related to research question 1 (effectiveness of the cross-age tutors on mathematical performance) as measured through the change in level, trend, and effect size on TEMI–AC weekly probe scores from baseline to intervention.

Research question 3: Tutor outcomes

The dependent variable aligned to the third research question was the classroom behaviors of students with EBD during their instructional class periods, as measured by Check-in/Check-out behavioral point sheets (CICO). The classroom behaviors were defined and aligned with the school- and class-wide behavioral expectations defined by the school in accordance with Positive Behavioral Interventions and Supports guidelines (PBIS; Horner, Sugai, Todd, & Lewis-Palmer, 2005; Sugai & Horner, 2006). Examples of these behavioral indicators fall under three overarching behavioral expectation categories: 'Be safe' (e.g., The student kept hands to self and did not touch classmates or their property without permission.), 'Be respectful' (e.g., The student was respectful of other students' feelings and avoided teasing them.), and 'Be responsible' (e.g., The student complied with adult requests without argument or complaint.). The tutors' daily classroom behavioral progress was measured through points obtained on the CICO, which contains a rating scale aligned to these expected behaviors (i.e., positively defined/stated). The CICO were scored by the students' classroom teachers who were blind to the study (i.e., paraprofessional for special education settings and general education teachers for inclusion classes) after each instructional class period and the data was evaluated through the overall points obtained (averaged by week), as well as the disaggregated data for average points obtained on tutoring (i.e., attended tutoring session) and non-tutoring days (i.e., did not attend tutoring session).

Research question 4: Maintenance

Two and four weeks after attending the last tutoring session, maintenance was assessed for tutees' mathematics performance on TEMI–AC measures. Similarly, tutors' classroom behavioral ratings on CICO point sheets were collected at the same rate as during the intervention phase (i.e., daily) to evaluate the maintenance of behavioral scores post-intervention (i.e., 2 and 4 weeks after last tutoring session).

Research question 5: Generalization

Tutees' generalization of early numeracy knowledge and skills was assessed through the administration of the Test of Early Mathematics Ability-3 (TEMA-3; Ginsburg & Baroody, 2003) as a distal measure, post-intervention. Similarly, tutors were assessed through a re-administration of the screening measure, Student Risk Screening Scale (SRSS; Drummond, 1994), which has been shown to be effective in determining changes students' risk for EBD over time (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008).

Research question 6: Social validity

All participating teachers, tutors, and tutees completed a social validity survey at the end of the study. Social validity measures were administered to assess perceived outcomes/effectiveness, as well as the practicality, feasibility, strengths and challenges of the intervention program. The survey contained both rating-scale and open-ended questions.

Measures

Screening for Children with Mathematics Difficulties and Distal Measures

Inclusion and exclusion criteria for tutees

First, the kindergarten teacher was asked to nominate students based on perceived difficulties in early numeracy knowledge and skills. These students were then screened for inclusion utilizing the Test of Early Mathematics Ability-3 (TEMA-3; Ginsburg & Baroody, 2003). Students met inclusion criteria for participation as tutees if their mathematics performance on the TEMA-3 ranked at or below the 25th percentile and therefore were considered at-risk for learning disabilities in mathematics and requiring intervention (Bryant et al., 2011; Fuchs et al., 2007). Students were excluded if they had an IEP, an Individualized Family Service Plan (IFSP), or were currently receiving instructional time or support for mathematics that was supplemental to their typical, business as usual (BAU) instructional minutes.

Test of Early Mathematics Ability-3 – Distal Measure

The TEMA-3 (Ginsburg & Baroody, 2003), which was used to screen tutees, was also utilized as a post-intervention distal measure. The TEMA-3 is a norm-referenced, diagnostic tool for determining mathematical strengths and weaknesses of students, ages 3 through 8, and consists of 72 items in the domains of informal and formal mathematics. Informal items evaluate four domains: numbering skills, number-comparison facility, calculation skills, and understanding of concepts. Formal items evaluate numeral literacy, mastery of number facts, calculation skills, and understanding of concepts. The assessment items frequently use representations in verbal, pictorial, and written formats. The reliability coefficients are reported to range from r = .94 to .96, and alternate-form coefficients range from r = .93 to .97. The test-retest reliability coefficients for TEMA-3 are r = .82 to .93. Criterion validity for TEMA-3 in relation to other standardized norm-referenced assessments of mathematics achievement show correlation coefficients ranging from r = .54 to r = .91 (Ginsburg & Baroody, 2003). Strong item validity is provided through correlations between individual items scores and the total scale score (Ginsburg & Baroody, 2003). Students whose TEMA-3 scores ranked at or below the 25th percentile qualified to participate in the study as tutees.

Screening for Students with EBD

Inclusion and exclusion criteria for tutors

Potential tutors were identified as fifth- and sixth-grade students who have or are at-risk for emotional-behavioral disorders (EBD). These students were initially identified through the investigator's assessment of students' IEP disability labels and IEP goals (i.e., related to/containing specific social skills or emotional-behavioral goals). Additionally, the special education teacher/case-manager of these fifth- and sixth-graders nominated students who were perceived to have the greatest challenges in problem classroom behavior in special education and inclusion instructional settings. All students who had IEP emotional-behavioral disabilities labels, goals, and/or were nominated by their special education teacher were screened individually, utilizing the standardized, norm-referenced Student Risk Screening Scale (SRSS; Drummond, 1994) to determine if they qualify as being at-risk for EBD. The special education teacher completed the SRSS for each of the students who initially qualified. Students qualified to be cross-age tutors if their SRSS score fell within the 'moderate' or 'high-risk category' (i.e., requiring supports or intervention). Additionally, the tutor's attendance was assessed prior to the start of the intervention to determine if they meet the requirement of consistent attendance rates. This was evaluated through attendance records from the previous year. If the student's daily attendance rate fell below 90% of the total number of school days for the previous year, the student was excluded from participating as a tutor. This exclusion criterion was included to ensure that tutees would not miss tutoring sessions due to a tutor's absence. Students who met inclusion criteria were then asked if they would be interested in participating in the program as tutors.

The Student Risk Screening Scale (SRSS; Drummond, 1994) is an evaluation measure developed to identify elementary students at-risk of anti-social, behavior problems (Drummond, 1994). In addition to be utilized as a screening measure for this study, the SRSS is also utilized as a post-intervention distal measure due to its ability to function as a progress monitoring measure over time, in the evaluation of students' risk level, through the aggregated data, for problem and antisocial behavior (Lane, Kalberg, Bruhn, Mahoney, & Driscoll, 2008). The screening measure consists of seven items rated on a 4-point Likert-type scale: *never* = 0, *occasionally* = 1, *sometimes* = 2, *frequently* = 3. The items were as follow: (1) steal; (2) lie, cheat, sneak; (3) behavior problem; (4) peer rejection; (5) low academic achievement; (6) negative attitude; and (7) aggressive behavior. Totals are then summed for all of the items and the student's total score is evaluated based on the following three categories of risk: low (0-3), moderate (4-8), or high (9-21). The measure suggests that students whose scores fall within the 'moderate' or 'high risk' categories should be provided supports or interventions for problem behaviors. The SRSS has been validated for use at the elementary, middle, and high school levels (Drummond, Eddy, & Reid, 1998; Lane, Kalberg, Parks, & Carter, 2008; Lane, Oakes, Ennis, Cox, Schatschneider, & Lambert, 2013; Lane, Parks, Kalberg, & Carter, 2007). Furthermore, this measure has been found to be psychometrically sound and socially valid in identifying students with externalizing and internalizing behaviors

(Lane, Bruhn, Eisner, & Kalberg, 2010; Lane, Little, et al., 2009). The special educator (i.e., case manager) for each of the students completed the SRSS and students whose SRSS score fell within the 'high-risk category' qualified to participate in the study as cross-age tutors.

Fidelity of assessment

A trained researcher with a background in special education, and familiar with the TEMA–3 procedures, observed, at minimum, 25% of the administrations of the TEMA–3 to prospective tutees. Fidelity of implementation of the assessment procedures was measured through a checklist aligned to the scripted prompts and procedures of the TEMA–3. Fidelity of assessment was calculated across administrations to average 97% (range = 94% to 100%).

Weekly Probes – Proximal Measures

Tutees: Texas Early Mathematics Inventory–Aim Checks

The Texas Early Mathematics Inventory–Aim Checks (TEMI-AC; University of Texas System/Texas Education Agency, 2009) are a researcher designed and validated early numeracy measures containing four subtests (i.e., magnitude comparisons, number identification, number sequences, quantity recognition), each taking 2-min to complete. This assessment measures numerical and operational skill and knowledge that are directly related to critical numerical competency and early mathematics skills (NCTM, 2008). The raw scores for the four subtests are summed which provided a total score. The TEMI-AC alternate-form reliability, across five forms, is above .80. The TEMI-AC were administered directly following the last tutoring session of each week within the tutoring setting by the special education teacher. Tutors served as supports during the administration by keeping their tutees on-task during the 2-min tests. Data was graphed weekly for each of the tutees by the investigator.

Tutors: Check-in/Check-out Behavioral Point Sheets

The Check-in/Check-out point sheet (CICO) is a behavior-rating scale completed by the classroom teacher of the student with EBD (i.e., cross-age tutor) after each class/period taught. CICO have been shown to be effective as progress monitoring measure for behavior in behavioral interventions (Campbell & Anderson, 2011; Dart et al., 2015). This type of measure can be found in similar incarnations, although containing differing indicators and other names, such as Daily Behavioral Report Card (e.g., Vannest, Davis, Davis, Mason, & Burke, 2010), but they usually contain the similar components (e.g., behavioral indicators) and similarly aligned rating scales for assessing in-class student behavior. Research suggests that implementers focus on 3-5 behavioral indicators and a measurement frequency that is consistent but manageable (Burke & Vannest, 2008). CICO scores are provided for target behaviors on a scale of zero to two points. The behaviors selected for measurement may be related to classroom and/or school-wide expectations, for example, "Be safe: keeping hands and feet to self". For this study, a score of two points indicates that the teacher was not required to redirect or warn the student about the behavioral indicator (e.g., keep hands to self) during the period. A score of one point indicates that the teacher was required to redirect or warn the student about a target behavior twice during the class period. And a score of zero points indicates that the teacher was required to redirect or warn the student about a target behavior two or more times during the given class period. The externalizing targeted behaviors for each student fell under the three school-wide Positive Behavioral Interventions and Supports (PBIS; Horner, Sugai, Todd, & Lewis-Palmer, 2005; Sugai & Horner, 2006) categories within the behavioral expectations matrix: 'be respectful', 'be responsible', and 'be safe'. Each of these behavioral categories contains specific school-wide PBIS behavioral expectations that have been defined and taught to the students through PBIS classroom lessons (e.g., Gelbar, Jaffery, Stein, & Cymbala, 2015; Barrett, Bradshaw, & Lewis-Palmer, 2008). Example student behaviors for each of the overarching expectation categories are below:

- Be Respectful: The student was respectful of other students' feelings and avoided teasing them.
- Be Responsible: The student complied with adult requests without argument or complaint.
- Be Safe: The student kept hands to self and did not touch classmates or their property without permission.

After each instructional period, the student's teacher (i.e., paraprofessional for special education settings and general education teacher for inclusion classes) scored their CICO point sheet. This allowed for behavioral ratings to be obtained from teachers who were blind or semi-blind to the study, supporting the validity of findings on the CICO measure in regards to generalized behavioral changes. One procedural option occasionally utilized in conjunction with the CICO behavioral progress monitoring measure is the check-in/check-out step between the student and teacher after each class period or at the start and end of the school day (e.g., Dart et al., 2015). During these brief meetings, the teacher and the student discuss why each individual score was given and discuss future goals and/or provide specific positive reinforcement. This step was intentionally omitted in this study's procedures, as supplemental one-on-one meetings between the tutor and the teacher would be considered an additional variable to account for, and thus, may influence internal validity.

Scoring and Interscorer agreement

The investigator trained the kindergarten teacher in a 30-min session on scoring procedures for the TEMI-AC. The initial interscorer agreement was assessed between the kindergarten teacher's and investigator's scoring of dummy coded, practice TEMI-AC forms and was calculated to be 97.8% across all forms. Interscorer agreement was calculated by summing the total number of agreements on participant responses/items and dividing by the total number of agreements plus disagreements and multiplying by 100 (Cihak & Bowlin, 2009; Haydon et al., 2012). The investigator scored all TEMI-AC tests during baseline, intervention, and maintenance phases, with kindergarten teacher also scoring 20% of the tests each week to determine interscorer agreement throughout the study. The average interscorer agreement rate was found to be 98.3%.

Teachers of the cross-age tutors (i.e., general education teachers for inclusion classes and a paraprofessional for special education classes) were trained on the CICO scoring procedures during a one period (45 m) training session and reliability was assessed through practice scenarios in the natural instructional environment through scoring the period following the training session. Additionally, reliability checks were conducted throughout the study for the teachers of the tutors, with a researcher attending and observing class periods containing all of the tutors and independently scoring a CICO for each. These reliability checks were conducted every other week during baseline and intervention phases. Interobserver/scorer agreement was calculated by totaling the number of agreements between the teacher and the researcher scores for each behavioral indicator for each student, dividing by the total number of comparisons (i.e., items) and multiplying that number by 100. Interobserver/scorer agreement ranged from 81.8% to 100% across all observed class periods (M = 89.6%).

MATERIALS

Number Line Board Games

Number line board games were selected as the intervention materials for this study for two reasons. First, the evidence base has shown that playing number line board games for as little time as 90-min over two weeks can be effective in promoting early numeracy knowledge and skills in verbal counting, numerical magnitudes, number comparison, and number line estimation for students from low-income backgrounds with similar academic needs to the participants in this study (Siegler, 2009; Siegler & Ramani, 2008). Second, the numerical board game's simplicity allows for tutors with basic mathematical knowledge and a familiarity with board game procedures (e.g., taking turns, spinning a spinner/rolling dice) to implement with minimal training time (e.g., Ramani, Siegler, & Hitti, 2012), and may be more appropriate for cross-age tutors, compared to the requirements of scripted curriculums or lesson planning (e.g., Blake, Wang, Cartledge, & Gardner 2000; Maher (1982; 1984). The tutor training procedures used in this study followed those used by Ramani, Siegler, and Hitti (2012) to train paraprofessionals to implement number line board games, although the language was adapted with the authors' permission, to meet the needs of the elementary student tutors being trained. The number line board games followed the specifications and designs used by Laski and Siegler (2014), where spaces contained numerals 0-100. The rationale for the selection of this range of numerals is based on the kindergarten CCSS in the area of number and operations (CCSS Initiative, 2010), which also aligns with the tutees' demographics and mathematics needs. The spinners that were used during the tutoring sessions contained the numerals 1, 2, and 3, and had an arrow affixed to the middle of the board.

PROCEDURES

Teacher Training

Training sessions took place during two, 45-min teacher planning periods over two consecutive days. Components of the training sessions included the administration procedures and use of progress monitoring measures for their individual students. For the kindergarten teacher, she was trained on the administration and scoring procedures for the tutees' weekly progress monitoring measure (TEMI-AC). The teachers of the tutors (i.e., general education teachers for inclusion classes and a paraprofessional for special education classes) were trained on the scoring procedures for the tutors' daily behavioral progress monitoring measure (i.e., CICO). Both the kindergarten teacher and the special education teacher of the tutors were trained on tutoring session supervision roles and responsibilities. Because the special education teacher of the tutors was already familiar with these students and their individual behavioral needs, he was provided the responsibility of being the lead superior of the tutoring sessions. These responsibilities included monitoring individual tutoring dyads, supporting tutor behavior, providing positive reinforcement to dyads, and managing the time of the sessions. This allowed for the kindergarten teacher to take on the responsibility of observing the fidelity of implementation of the tutoring procedures by the cross-age tutors through the use of the fidelity observational checklist that the teacher was provided training on during the teacher sessions. The investigator conducted interobserver agreement checks and follow-up training sessions with the individual teachers at scheduled times throughout the intervention. Scheduling of tutoring session times were mutually agreed upon by the two participating teachers prior to the start of the study. It was determined that the 'exploratory period' of the day when the kindergarten classroom conducted learning centers/stations would be the most appropriate time for tutoring sessions and that these sessions would take place during the same times and days during the school week.

Baseline Phase

During the baseline phase, tutors and tutees attended their business as usual (BAU) class schedules. The TEMI-AC was administered to the tutees weekly at approximately the same time of the school day when future tutoring sessions would be implemented. Tutors' classroom behaviors were also continually assessed through the scoring of CICO sheets (i.e., by general and special education teachers) after each class period. Once a stable baseline (i.e., level, trend, variability; Horner et al., 2005; Kennedy, 2005; Kratochwill et al., 2010; Parsonson & Bear, 1978) was determined to be present for a given tutor/tutee dyad, based on the tutee's performance on the weekly TEMI-AC and the tutors' scores on CICO point sheets, a 1:1 tutor training session was provided to the tutor and the intervention tutoring sessions commenced the following day.

Intervention Phase

Tutor training

After a stable baseline was identified for a tutor/tutee dyad, and prior to the start of attending tutoring sessions, each tutor was individually trained on tutoring procedures by the investigator during a 1:1, 45-min (i.e., one class period) training session. The tutor training sessions included introducing the number line board game materials, modeling, guided practice, corrective feedback, role-playing tutoring sessions, and evaluating the following skills: (1) instructional techniques and number line board game procedures, (2) corrective feedback methods, and (3) positive behavioral reinforcement strategies.

Instructional techniques included: how to greet their tutee, reviewing the previous session (i.e., what went well, what they will focus on improving during the current session), starting the game, keeping their tutee on task, and the number line board game rules and procedures. Training on the number line board game procedures and corrective feedback methods were based upon the training manual developed by Ramani, Siegler, and Hitti (2012) to train paraprofessional for an earlier study. Some of the language used in training manual was adapted, with the authors' permission, to meet the needs of the elementary student tutors. The corrective feedback method included a two-step process. When the tutee was observed making a counting error, for example, if the tutee's game piece sat on number six and they spun a three on the spinner and they then moved their game piece by counting the number of spaces they earned (e.g., "1, 2, 3") instead of counting-on from the number their game piece was currently sitting on (e.g., "7, 8, 9"), the tutor would first verbally prompt the tutee to count the numbers on the board game, giving them another practice opportunity. If the tutee again made an error by counting from one, or made an error in counting the consecutive numbers correctly (e.g., "7, 9, 10"), the tutor would then model by showing the tutee how to count-on correctly, followed by giving the tutee another opportunity to practice after the model. Positive behavioral reinforcement strategies included providing feedback on tutees' counting skills and general behavior through specific positive behavioral statements (i.e., praise for specific behaviors). Examples include: "I like how you took your time and counted-on from the number you started on."; "You are doing a great job sitting in your seat and staying focusing on the game today."

The quality of implementation of tutor training sessions was assessed through a procedural checklist containing the training components aligned to the scripted training manual developed by Ramani, Siegler, and Hitti (2012). A second researcher was trained on the observational measure and IOA was established at greater than 95% on practice administrations of the training prior to implementation. The second researcher observed 40% of the tutor training sessions and quality of implementation of tutor training procedures by the investigator ranged from 94%-99% across sessions (M = 97%; IOA = 96%).

Tutoring sessions

Tutees attended three tutoring sessions per week in the resource classroom over the duration of 10 weeks. Each session lasted approximately 25-30-min and was supplemental to the tutees' BAU daily instructional time in mathematics (i.e., 45-min per day). None of the tutees were receiving additional instructional time or support for mathematics at the time of intervention. These sessions took place on the same days and times each week. Tutees participated in the number line board game with the tutor, one-on-one, with the tutor providing the game procedures, modeling, corrective feedback, and positive reinforcement. After the last session for the week, tutees were administered a paper-based progress monitoring check (TEMI-AC), which consists of four subtests, each 2-min in length.

Tutors' Fidelity of Implementation

The kindergarten teacher assessed fidelity of implementation of tutoring procedures by the cross-age tutors with EBD during each tutoring session. The teacher, trained on the fidelity checklist of tutoring procedures (i.e., modeling, providing practice opportunities, corrective feedback procedures, positive behavioral reinforcement techniques), observed using momentary time-sampling procedures, rotating among each tutor every 30 s. The teacher used a stopwatch or iPhone with a buzzer alarm to notify a change to the next interval. IOA was assessed through a second, independent observation of the same tutoring sessions by a trained researcher for more than 30% of the total fidelity observations conducted during the study. The number of item agreements between the two observers were summed and then divided by the total number of item agreements plus disagreements, and then multiplied by 100. The interobserver agreement ranged from 88% to 100% (M = 92.8%). When a tutor's fidelity of implementation of tutoring procedures was observed to fall below an 80% average across a given week, a retraining session on tutoring procedures was provided to the tutor by the investigator immediately prior to the first tutoring session the following week.

Tutor Retraining Sessions

Retraining sessions were 30-min in length and contained the same procedures as the initial tutor training. A majority of session's time was spent on modeling, guided practice, role-playing, and feedback on the tutoring components that were identified through the fidelity of implementation checklist as the most frequently omitted during the tutoring sessions. At the end of the retraining session, the tutor would role-play an abbreviated tutoring session with the investigator, where the investigator would take the role of the tutee and assess the tutor's fidelity of implementation of tutoring procedures using the same fidelity of implementation checklist. Additionally, when the school schedule experienced a break of four or more consecutive days, tutors were provided a retraining, booster lesson upon returning to school (i.e., prior to the next scheduled tutoring session).

Maintenance

The maintenance phase took place for four weeks after the conclusion of the last tutoring session. No further tutoring sessions took place between the end of the intervention phase and the administration of maintenance measures. To assess maintenance, the TEMI-AC was administered to each of tutees during the typically scheduled tutoring time two and four weeks after the final tutoring session. To assess the maintenance of outcomes for the tutors, teachers of the tutors was asked to continue scoring CICO point sheets daily for each of the tutors over a 4-week post-intervention phase.

Generalization

During this maintenance phase, tutees were also administered a post-intervention, distal measure (i.e., TEMA-3) to assess the generalization of early numeracy knowledge and skills. Additionally, the SRSS was administered as a generalization measure for each of the tutors, two weeks after tutoring sessions ended. The special education teacher completed this measure for each student to determine if risk status for EBD had changed from pre-intervention to post-intervention.

Social Validity

Each of the students, tutees, and tutors, completed a researcher-developed social validity questionnaire after the intervention. The teachers also provided responses related to observed changes in mathematics abilities of the tutees and classroom behaviors of the tutors. Questions included whether the students' performance in mathematics/classroom behaviors had changed due to the intervention, favorite aspects of the tutoring program, challenges of the tutoring program, and whether they would like to participate in future incarnations of the program. The survey recorded responses through rating scales and open-ended questions. Rating options were 1 = strongly disagree, 2 = disagree, 3= neutral/no change, 4 = agree, 5 = strongly agree. The tutor survey contained nine rating-scale questions, the teacher survey contained eight rating-scale questions, and the tutee survey contained seven rating-scale questions; and all surveys contained two open-ended questions. Due to the young age of the tutees, the researcher recorded the dictated responses of each student individually. The tutors and the teachers all recorded their responses in written form. Tutors were also provided the same supports when if their writing abilities were not proficient enough to provide full responses. This was determined by asking the teacher which of the tutors would benefit from this accommodation. All students were verbally prompted to try to provide additional information for responses, verbal or written, when necessary.

DATA ANALYSIS

Student data obtained on the measures outlined above was entered into a secure database at the end of each week. Raw data were secured in a locked cabinet in the researcher's office. All digital data from the assessments was stored on a secure server through the College of Education. Data were graphed and analyzed on a weekly basis for the TEMI-AC score and the CICO scores by the investigator to determine baseline and intervention progress by tutees and tutors, respectively.

Visual Analysis of Tutee and Tutor Proximal Measures

Visual analysis procedures were utilized to evaluate the effectiveness of the cross-age tutoring program on tutees' mathematical performance and tutors' classroom behaviors (Kennedy, 2005). This analytical procedure requires the assessment of participants' response on dependent variables through graphical data, across phases (i.e., baseline, intervention, maintenance; Parsonson & Baer, 1978). Visual data in graphical form is evaluated to determine if a casual relation between the independent variable and participant outcomes is present (Kennedy, 2005). Additionally, the visual analysis can determine the strength or evidence of the casual relation based on six features of the data: (1) level, (2) trend, (3) variability, (4) immediacy of change/effect, (5) overlap of data, and (6) data patterns consistency across phases (Horner et al., 2005; Kennedy, 2005;

Kratochwill et al., 2010). Kratochwill and colleagues (2010) define level as the average of the scores across a given phase. The trend can be determined by fitting a straight line (i.e., 'trend line') through the phase's data to determine the slope (Kratochwill et al., 2010). Variability within the phase can be determined by way of the standard deviation of the data in relation to the trend line (Kratochwill et al., 2010). The immediacy of effect can be identified by evaluating the change in level between phases trough the assessment of the first and last three data points 'connecting' the two phases (Kratochwill et al., 2010). Overlap of data is the percentage of data from a given phase that overlaps with the data in the following phase (Kratochwill et al., 2010). Finally, identifying the consistency of data requires the assessment of data across similar phases, and identifying if the patterns across the phases are similar from one to the next (Kratochwill et al., 2010). In sum, visually evaluating these characteristics of the graphical data allows researchers to identify effects of an independent variable on dependent variables (Kratochwill et al., 2010). Tutee and tutor data points from baseline through intervention phases, as well as within the maintenance phase (i.e., 2 and 4 weeks after attending the last tutoring session) were assessed utilizing theses visual analysis procedures. Additionally, effect sizes were calculated and interpreted to evaluate the overall effects of the intervention program for both participant populations.

Proximal Effect Sizes

For single-subject designs, the percentage of data points exceeding the median of the baseline phase (PEM) approach was chosen to assess the effectiveness of outcomes, due to its assumed validity in assessing disruptive behaviors (Chen & Ma, 2007), a frequently targeted skill for students with EBD. Additionally, when floor or ceiling data points are present PEM is still capable of reflecting effect size. PEM is calculated by identifying the median baseline point and drawing a median line from that point through intervention phases. The percentage of data points above or below the median line is calculated by summing all intervention data points above or below (depending on targeted skill or measure; e.g., increasing an academic skill or decreasing a behavior) the line and dividing that sum by the total number of data points in the intervention phase. PEM results were analyzed using the following scale: 90-100% = large or highly effective, 70%-90% = moderately effective, and < 70% = small or questionable effectiveness (Ma, 2006).

Generalization

The generalization of behavioral change of cross-age tutors with EBD was assessed through a post-intervention administration of the SRSS, completed by the special education teacher of the tutors 2- and 4-weeks after attending the last tutoring session. Generalization was analyzed by identifying changes in tutors' risk for EBD. An overall effect size was calculated by comparing pre- and post-intervention SRSS raw scores, across tutors. Generalization of tutees' early numeracy knowledge and skills was also evaluated. Two-weeks after attending the last tutoring session, the TEMA–3 was administered to the tutees. The scores obtained on this measure were compared to the pre-intervention scores obtained on the same measure and evaluated using the effect size calculation and analysis procedures outlined below.

Distal Effect sizes

To aid in interpreting the results, Hedges' *g* (Hedges, 1981) effect sizes were calculated for distal measures administered to both tutees and tutors. Hedges' *g* was chosen due to its ability to represent individual level effect size and provide a better estimate for small samples. Additionally, What Works Clearinghouse supports Hedges' *g* as an effective method in analyzing an intervention's effectiveness (WWC, 2014). Hedges' *g* was calculated using the following formula:

$$g = M_1 - M_2 / S_{pooled} \text{ where } S = \sqrt{[\sum(X - M)^2 / N\text{-}1]} \text{ and } S_{pooled} = \sqrt{MS_{within}}$$

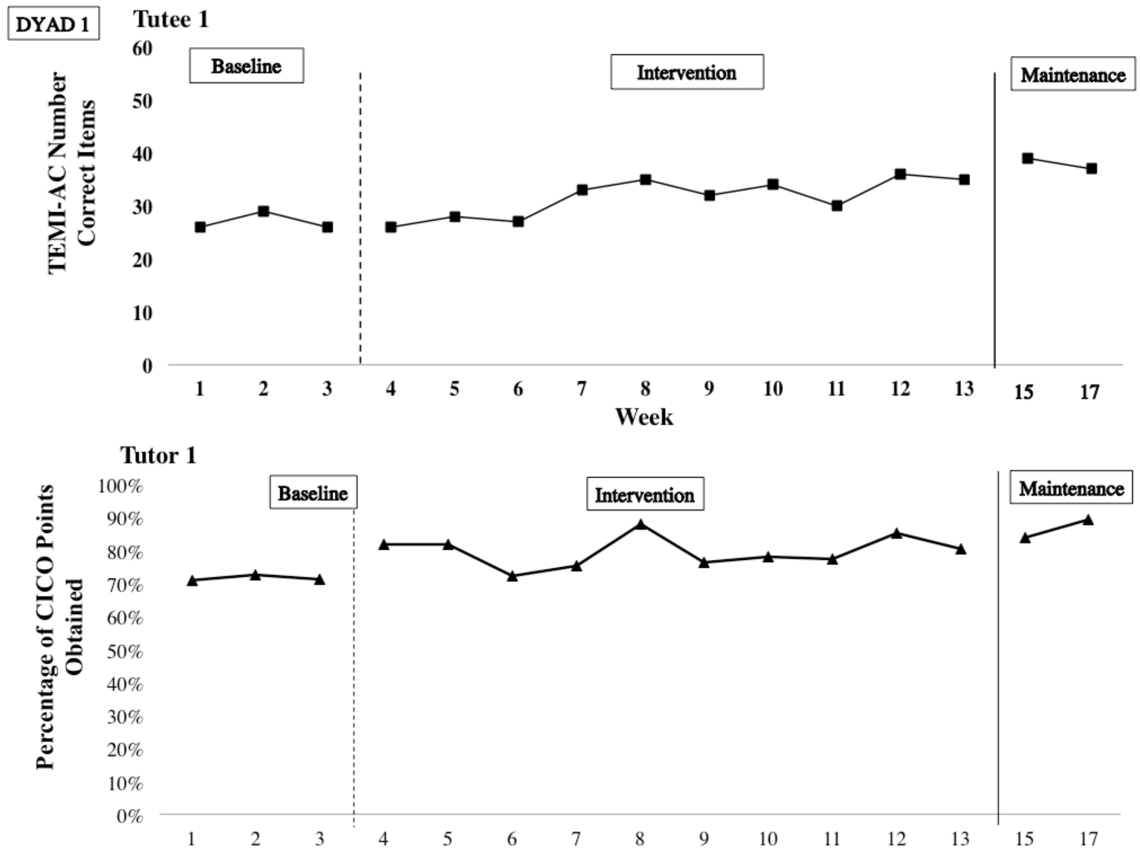
Effect sizes are interpreted based on the criteria set by Cohen (1969), where 0.2-0.49 is interpreted as a small effect. 0.5-0.79 shows medium or moderate effect, and 0.8 and greater being large or significant effect.

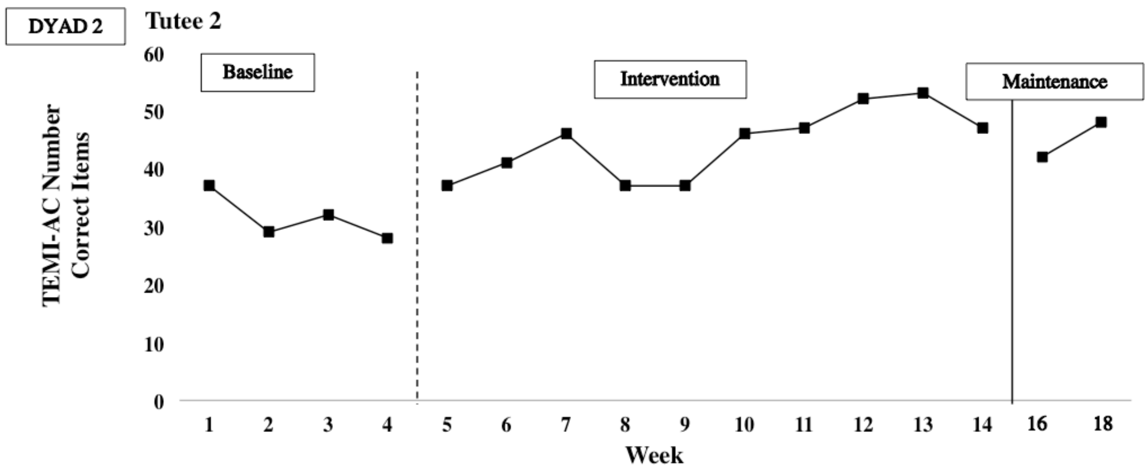
Chapter 4: Results

The purpose of this study was to investigate the potential effects of a cross-age tutoring model on the mathematics performance of young students at-risk for mathematics difficulties. Additionally, potential effects were also assessed for the tutors who were students identified with, or at-risk for EBD. The research questions for the intervention study were as follow:

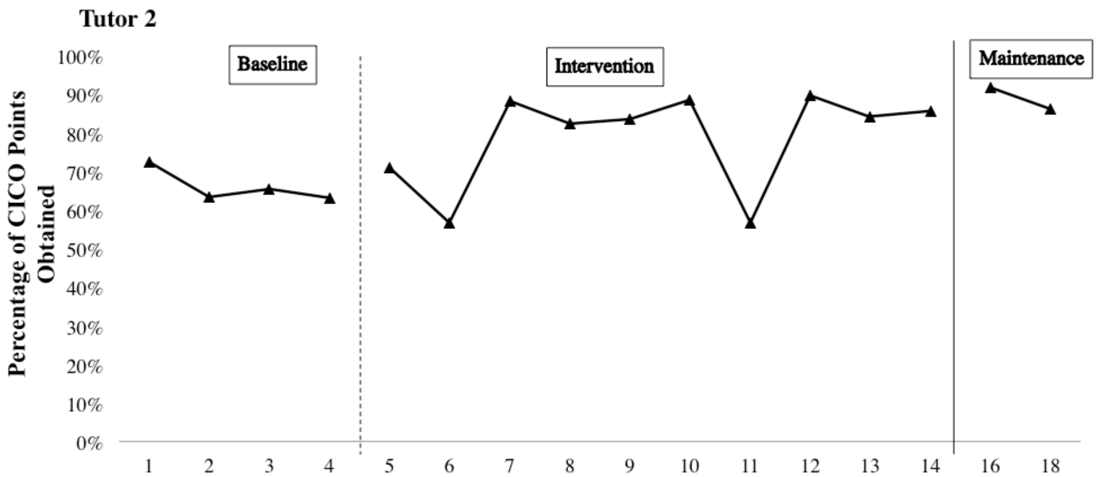
1. What effects will a number line game delivered by a cross-age tutor with EBD have on the early numeracy skills of kindergarten students at-risk for math disabilities?
2. Can students with EBD effectively serve in the role of cross-age tutors (i.e., implement instruction with fidelity and increase tutees' number sense skills)?
3. What effect(s) will the training and implementation of the cross-age tutoring program have on the tutors' behavioral ratings on CICO point sheets?

This chapter is organized in alignment with the research questions and their related results. Post-intervention, social validity survey results are also presented. Finally, it should be noted that the arrangement of tutors and tutees to dyads follows the same organization as previous chapters, where tutors and tutees are aligned by number (e.g., Tutee 1 and Tutor 1 formed dyad 1, Tutee 2 and Tutor 2 formed dyad 2). Figure 4.1 displays proximal data related to tutees' weekly total scores (i.e., TEMI-AC) and tutors' weekly average behavioral points obtained (i.e., CICO), arranged by dyad.

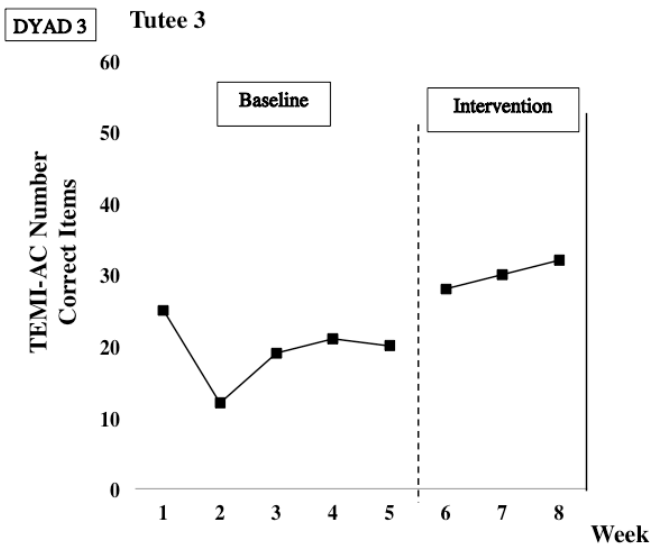

DYAD 1
Tutee 1
TEMI-AC Number Correct Items
Baseline
Intervention
Maintenance
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 15 17
Week
Tutor 1
Percentage of CICO Points Obtained
Baseline
Intervention
Maintenance
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 15 17

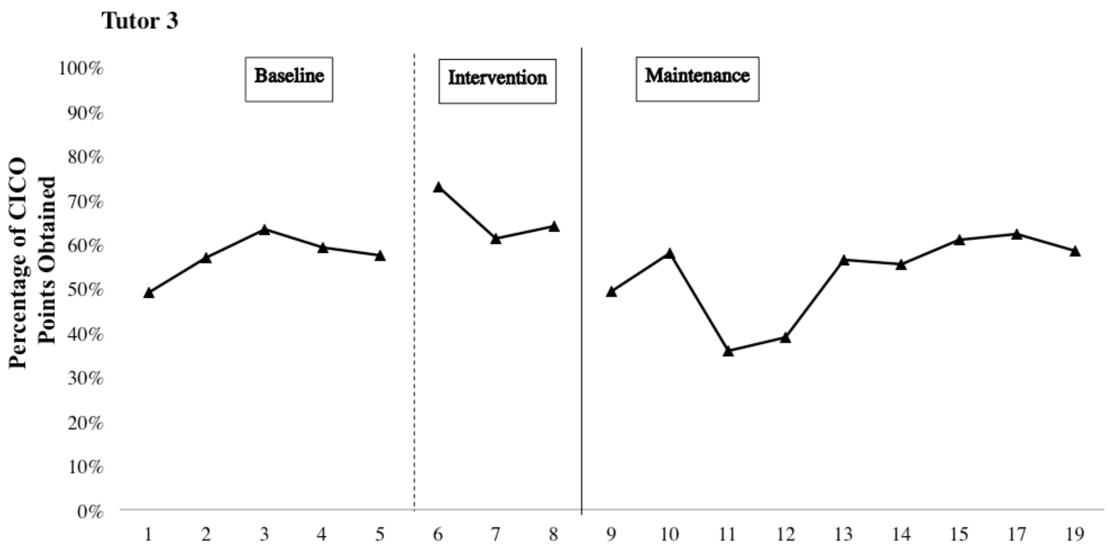


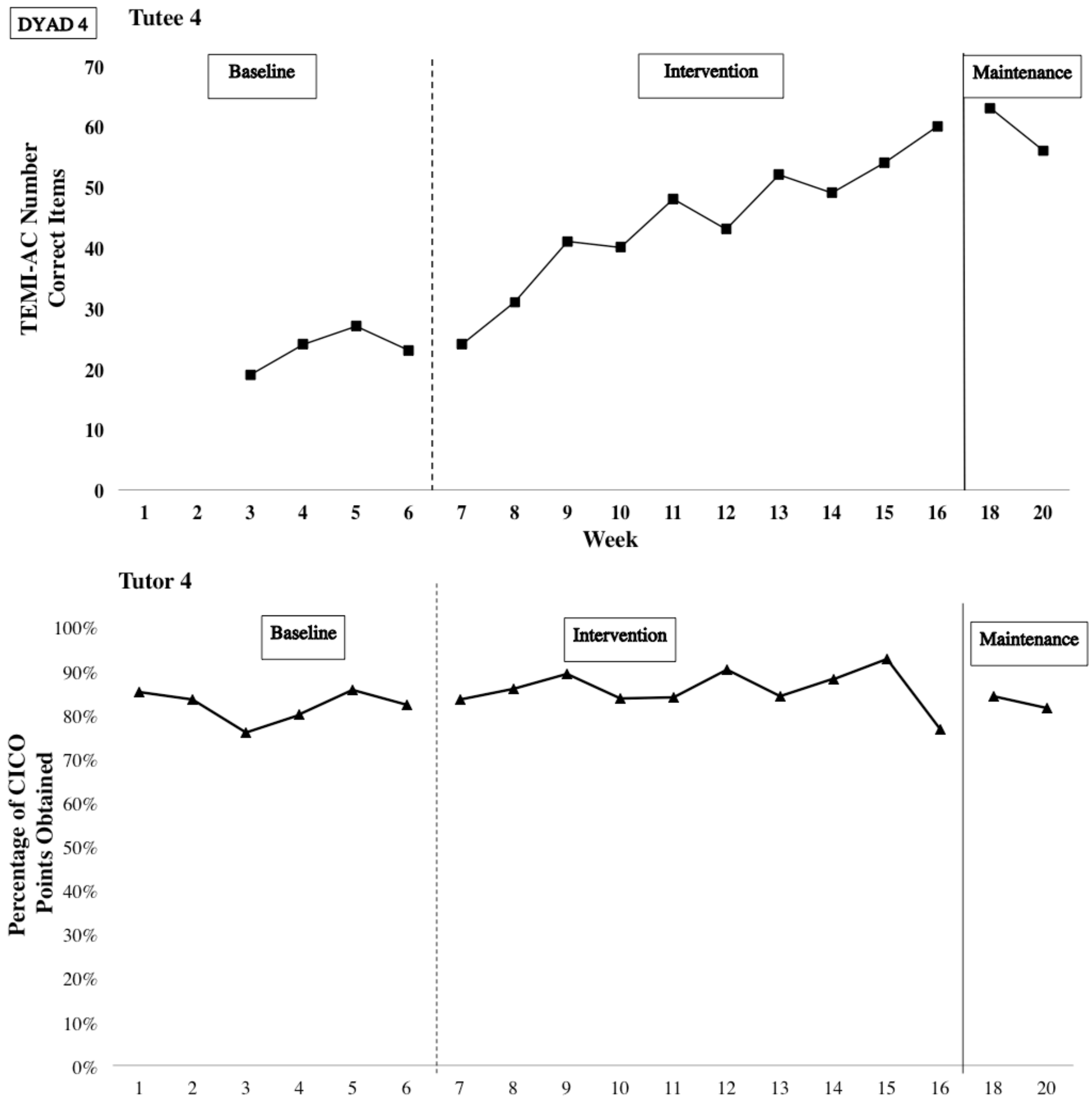
DYAD 2
Tutee 2
Baseline
Intervention
Maintenance
TEMI-AC Number Correct Items
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 18
Week

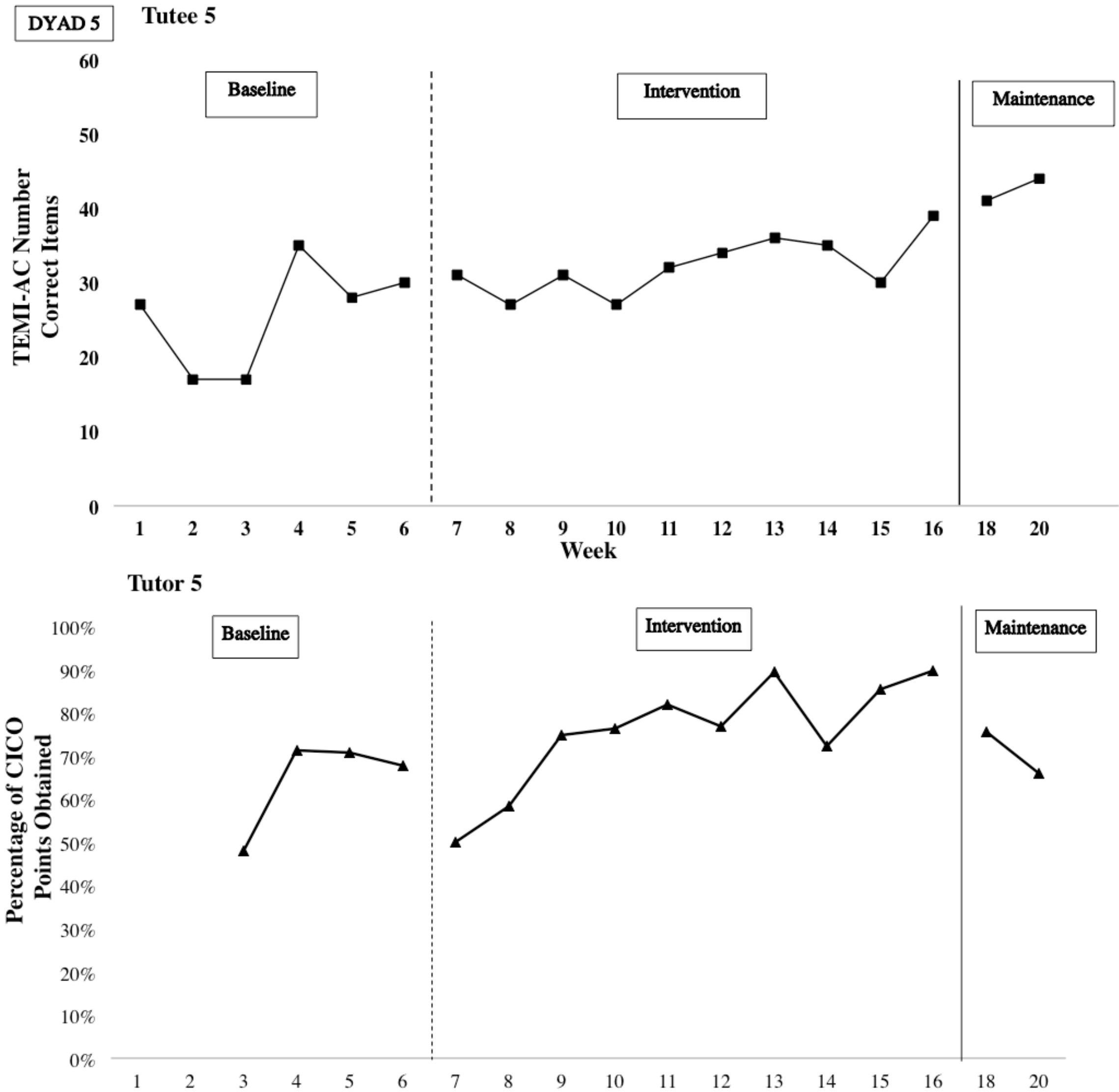


Tutor 2
Baseline
Intervention
Maintenance
Percentage of CICO Points Obtained
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 18


DYAD 3
Tutee 3
Baseline
Intervention
TEMI-AC Number Correct Items
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
Week


Tutor 3
Baseline
Intervention
Maintenance
Percentage of CICO Points Obtained
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
17
19


DYAD 4
Tutee 4
Baseline
Intervention
Maintenance
TEMI-AC Number Correct Items
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 20
Week
Tutor 4
Baseline
Intervention
Maintenance
Percentage of CICO Points Obtained
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 20



Note. TEMI-AC = Texas Early Mathematics Indicators–Aim Checks: total scores; CICO = Check-in/Check-out behavioral point sheet

Figure 4.1: Proximal Data for Tutees and Tutors Arranged by Dyad

RESEARCH QUESTION 1

Research question 1 examined the effects of a number line board game, delivered by a cross-age tutor with EBD, on the mathematics performance of kindergarten students at-risk for math disabilities. To assess the early numeracy knowledge and skills of these students at-risk for mathematics disabilities (i.e., tutees) weekly probes were administered in the form of the Texas Early Mathematics Inventory–Aim Checks (TEMI-AC; University of Texas System/Texas Education Agency, 2009). The four subtest scores were combined to create a total score, which was then used to evaluate students' progress on a weekly basis.

Visual Analysis

As recommended by What Works Clearinghouse, the six features of visual analysis of single-case data include the evaluation of level, trend, variability, immediacy of effect, overlap, and consistency of data patterns (Kratochwill et al., 2010). These features were analyzed to determine if a casual relation existed between the number line board game, cross-age tutoring program (i.e., independent variable) and the tutees' early numeracy knowledge and skills as exhibited through their performance (i.e., total score) on weekly TEMI-AC probes (i.e., dependent variable). Figure 4.2 displays tutees' weekly TEMI-AC total scores during each phase. Figures 4.3 and 4.4 show level and trend data, respectively, along with disaggregated TEMI-AC subtest scores. Table 4.1 shows variability, immediacy of effect, and overlap data for tutees.

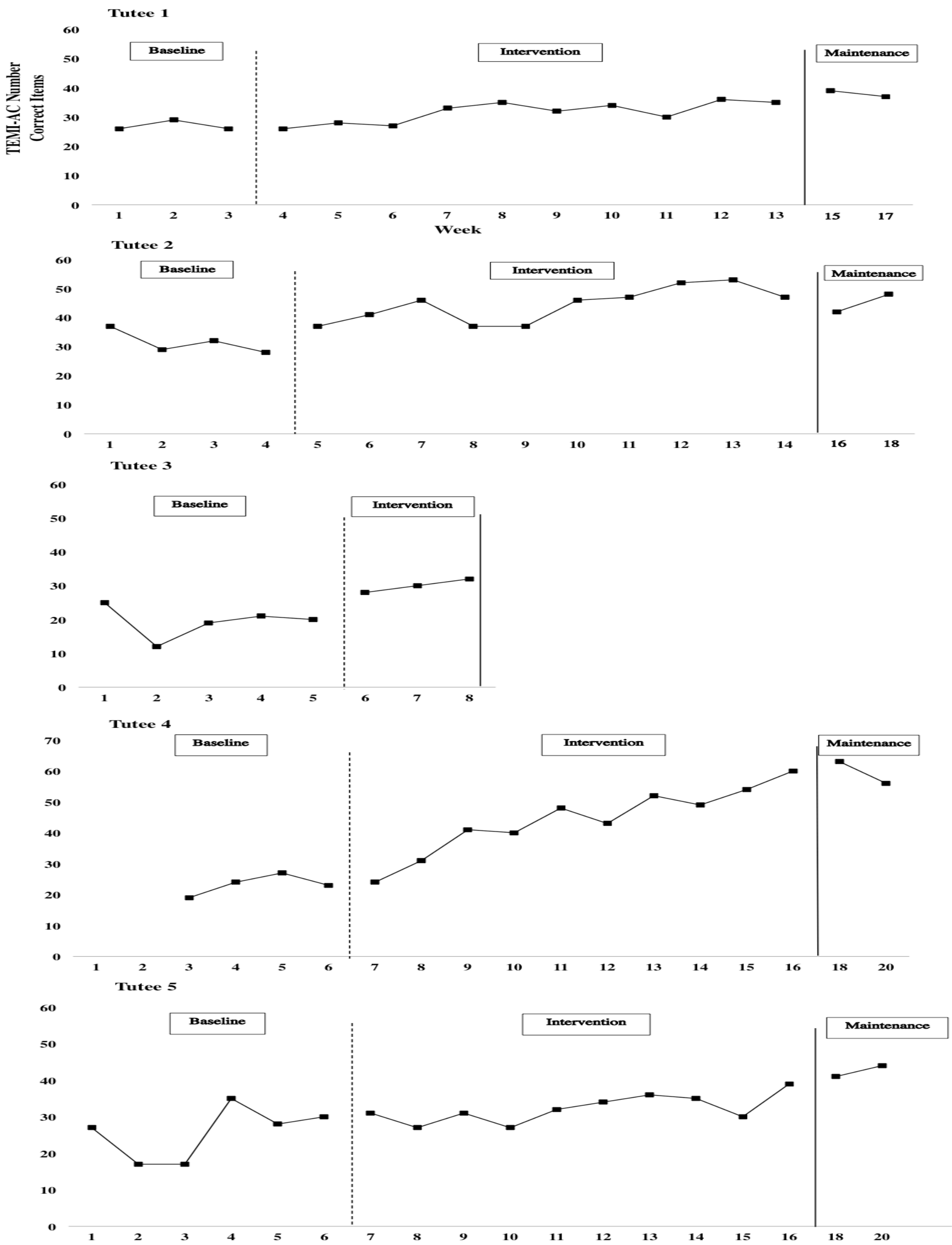
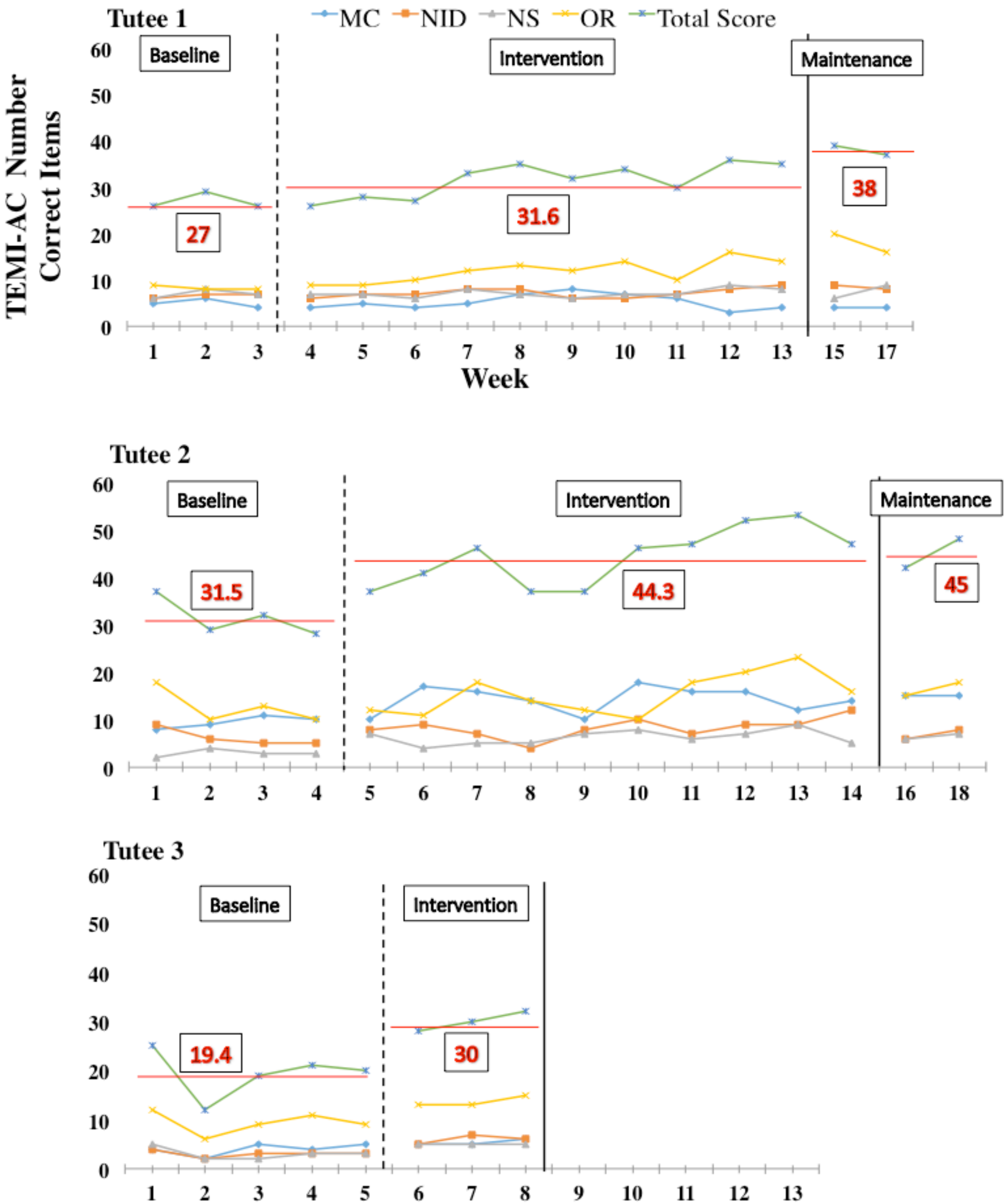
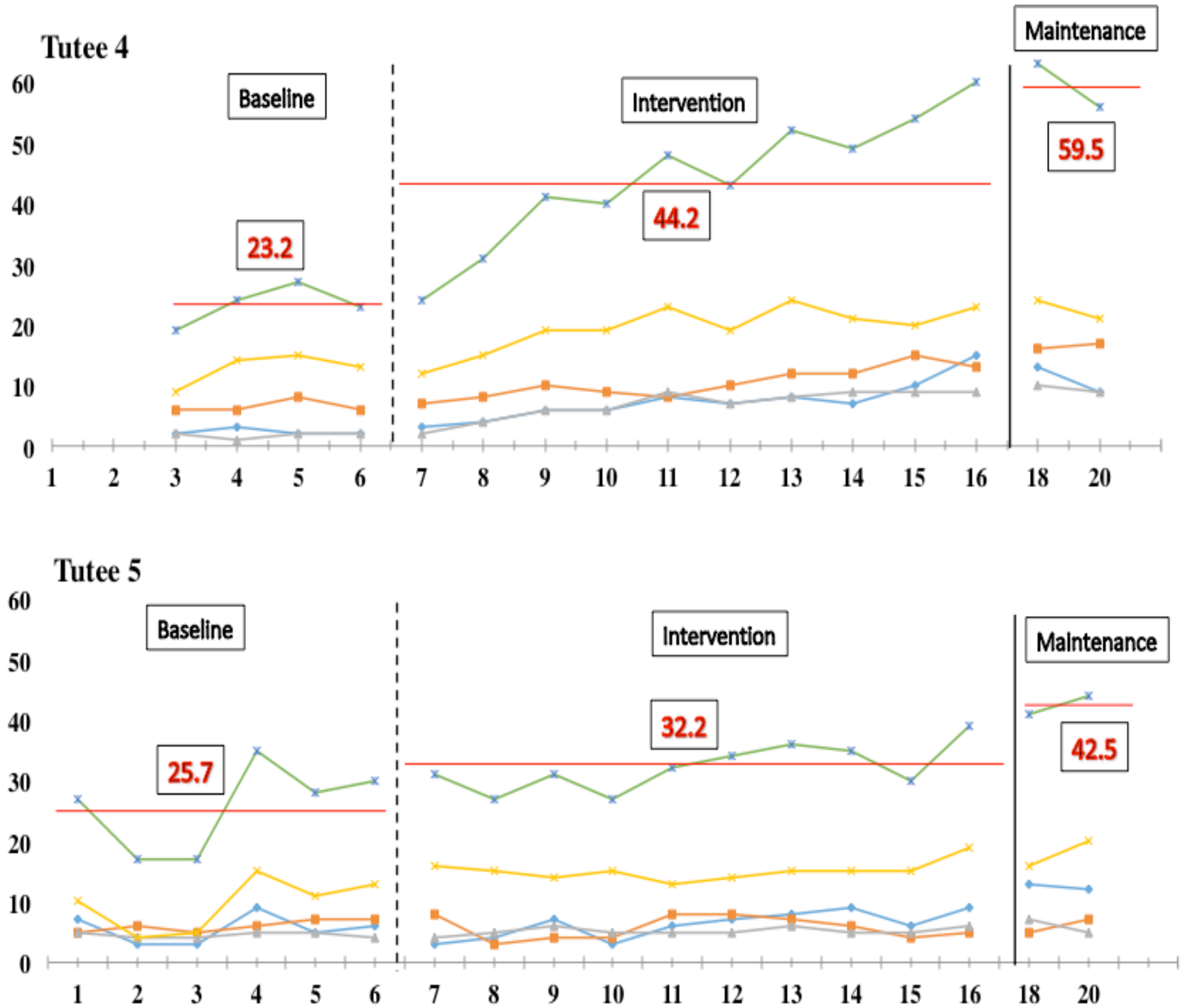


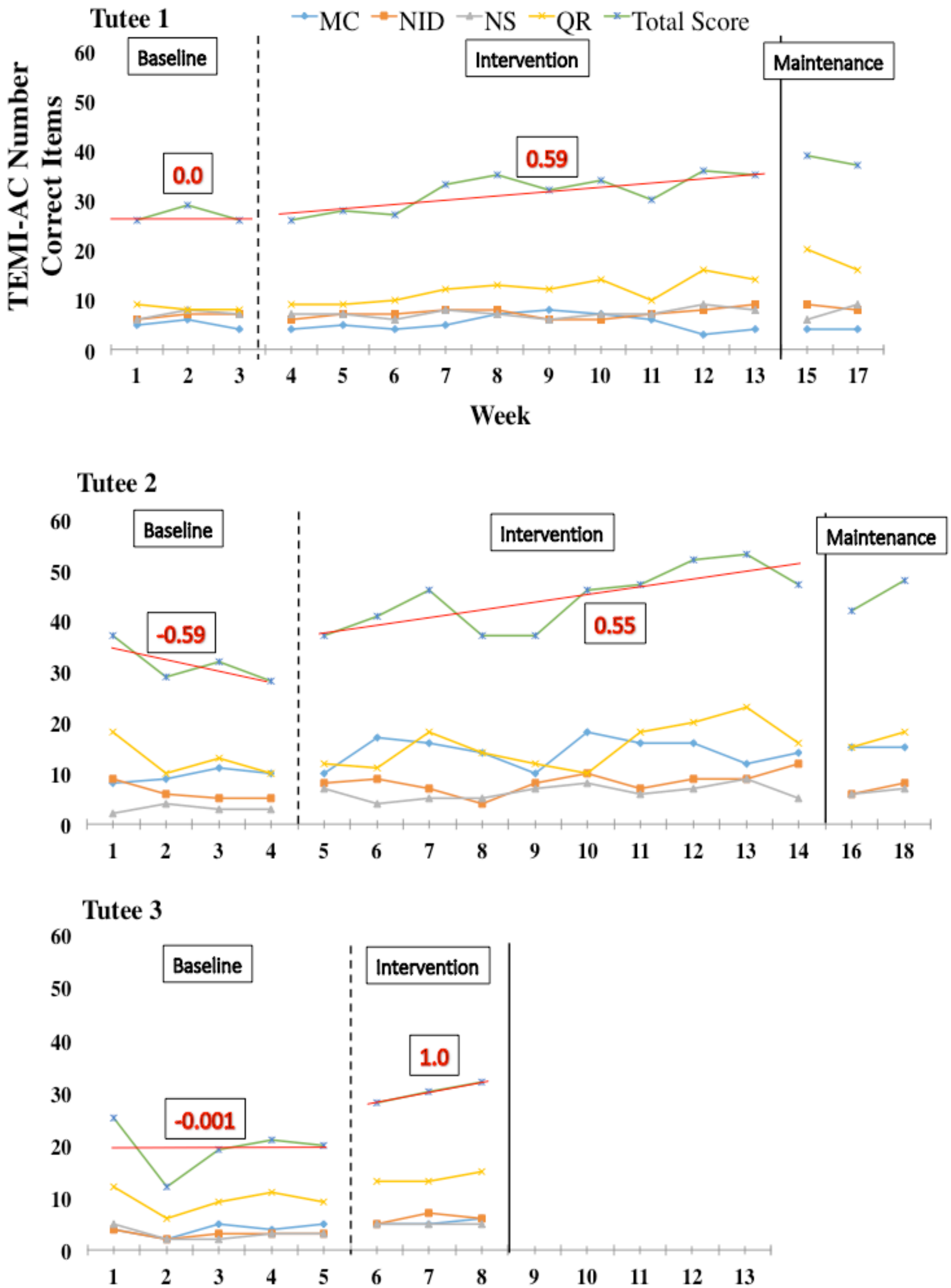
Figure 4.2: Tutees' TEMI-AC Total Scores by Week


Tutee 1
MC
NID
NS
OR
Total Score
TEMI-AC Number Correct Items
Baseline
Intervention
Maintenance
27
31.6
38
Week
Tutee 2
Baseline
Intervention
Maintenance
31.5
44.3
45
Tutee 3
Baseline
Intervention
19.4
30

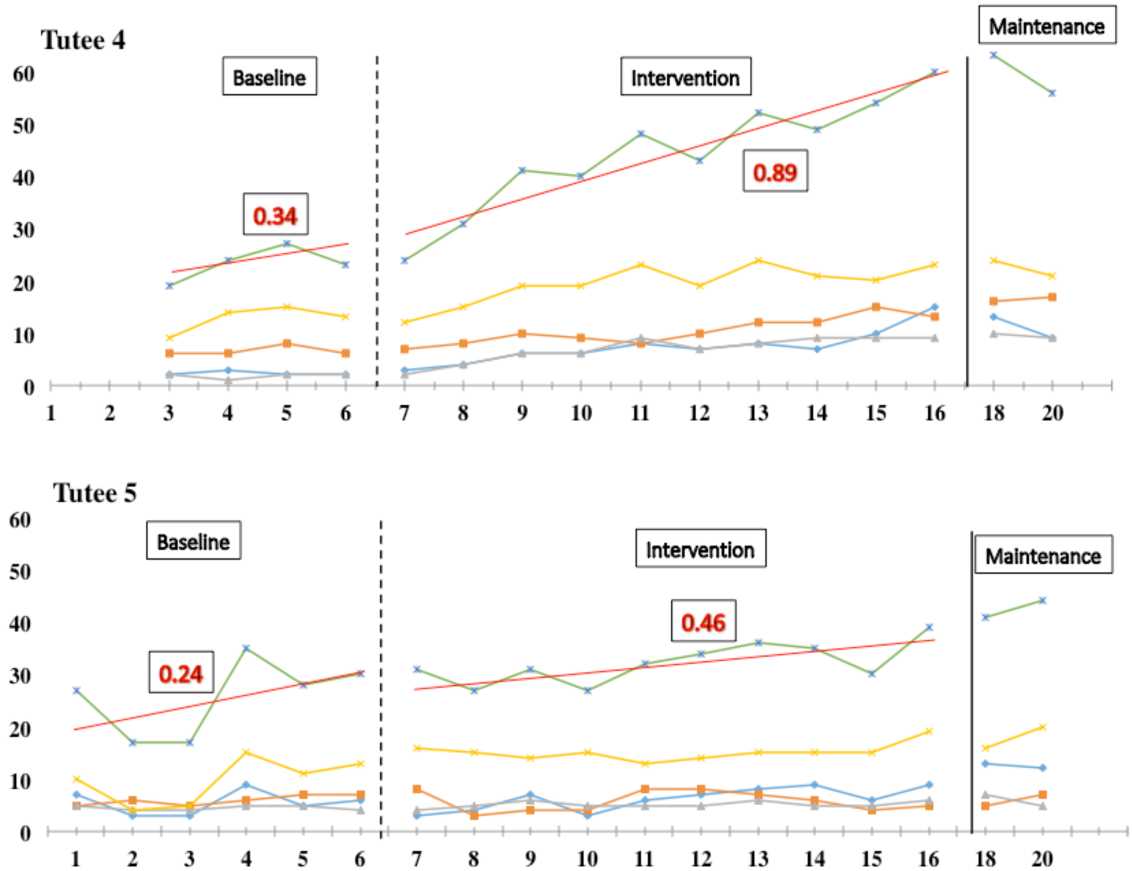


Note. MC = magnitude comparison; NID = number identification; NS = number sequence; QR = quantity recognition

Figure 4.3: Level and Disaggregated Data for Tutees' Weekly TEMI-AC Scores



Tutee 1
MC
NID
NS
QR
Total Score
TEMI-AC Number Correct Items
Baseline
Intervention
Maintenance
0.0
0.59
Week
Tutee 2
Baseline
Intervention
Maintenance
-0.59
0.55
Tutee 3
Baseline
Intervention
-0.001
1.0



Note. MC = magnitude comparison; NID = number identification; NS = number sequence; QR = quantity recognition

Figure 4.4: Trend and Disaggregated Data for Tutees' Weekly TEMI-AC Scores

Note. SD = standard deviation

Tutees	Variability standard deviation (range)		Immediacy of effect (%)	Overlap
	Baseline	Intervention		
Tutee 1	1.73 (26-29)	3.63 (26-36)	0.0	Yes
Tutee 2	4.04 (28-37)	6.02 (37-53)	116.67	No
Tutee 3	4.72 (19-25)	2.00 (28-32)	100.0	No
Tutee 4	3.30 (19-27)	10.83 (24-60)	73.33	No
Tutee 5	7.26 (17-35)	3.85 (27-39)	-13.33	Yes
Mean (SD)	4.21 (2.04)	5.27 (3.42)	55.33 (58.86)	

Table 4.1: Tutees' Variability, Immediacy of Effect, and Overlap data.

Tutee 1

Tutee 1's weekly TEMI-AC total score level during baseline phase was 27.0 and then increased to 31.6 during intervention (see Figure 4.3, Panel 1). After the introduction of the intervention (i.e., the tutee began attending the weekly tutoring sessions), a change in level was not immediately present, as he was one of two tutees to show overlap of intervention and baseline data (see Table 4.1). This overlap of data between phases occurred within the first three weeks of the intervention and all remaining weeks of showed scores above baseline level. Tutee 1 showed the most stable baseline of all tutors, with no directional trend in TEMI-AC total scores (see Figure 4.4, Panel 1). During the intervention phase, the tutee showed an upward trend in TEMI-AC total scores (0.59). The level of maintenance data (38.0) was higher than the level during intervention phase showing Tutee 1's maintenance and continued his upward trend in performance on

TEMI-AC total scores two and four weeks after the intervention was completed. Evaluating the consistency of data patterns showed Tutee 1 to have least amount of variation in baseline data (SD = 1.73; range = 26-29) compared to other tutors and relatively low variation during the intervention phase (SD = 3.63; range = 26-36).

Tutee 2

Tutee 2's TEMI-AC total score level from baseline to intervention phase increased from 31.5 to 44.3 (See Figure 4.3, Panel 2). Change in level was observed as Tutee 2's first TEMI-AC total score during intervention was well above any of the last three baseline data points (116.67%; see Table 4.1). This was the largest immediacy of effect among all of the tutees. Overlap of data was not present for Tutee 2. Trend during baseline phase was downward (-0.59) but the last three data points before intervention were stable (See Figure 4.4, Panel 2). During intervention the trend in TEMI-AC total scores was upward (0.55). Both maintenance data level (45) was above baseline level (31.5) and were slightly higher than intervention phase level (44.3). Baseline fluctuation around mean score was relatively stable compared to other tutees (SD = 4.04; range = 28-37), and this level of variability continued within intervention phase scores (SD = 6.02; range = 37-53).

Tutee 3

Tutee 3 moved out of the district three weeks into the intervention phase, and therefore, a decision was made to revert his tutor, Tutor 3, back to baseline phase (i.e., remove tutoring program responsibilities). This decision was made for two reasons: first, to assess any change in data patterns after the removal of the intervention, and second, to determine any potential long-term maintenance of CICO behavioral scores for the tutor.

Tutee 3's baseline level was stable with no directional trend (-0.001; see Figure 4.4, Panel 3) and a level of 19.4 (see Figure 4.3, Panel 3). During the three weeks of intervention, the level increased to 30 and demonstrated an upward trend (1.0). His immediacy of effect (100%; see Table 4.1) was the second highest across tutees, and there was no overlap of data between phases. Variability during baseline was average compared to other tutees (SD = 4.72; range = 19-25) and continued to stabilize during intervention (SD = 2.00; range = 28-32). Maintenance data was unavailable for Tutee 3 due to his exiting of the district before the completion of the intervention.

Tutee 4

Tutee 4's weekly TEMI-AC total score level during baseline phase was 23.2 and then increased significantly, to 44.2, during intervention (see Figure 4.3, Panel 4)). A change in level was demonstrated (73.33%), but no overlap data was present across phases. Tutee 4 showed an upward trend in TEMI-AC total scores during baseline (0.34) and a continuing upward trend during intervention at an increased rate (0.89; see Figure 4.4, Panel 4). The data level of the maintenance phase (59.9.) was well above intervention phase level (44.2) and significantly higher than baseline level (23.2). Evaluating the consistency of data patterns showed Tutee 4 to have average variation in baseline data (SD = 3.30; range = 19-27) compared to other tutees and significantly high variation during the intervention phase (SD = 10.83; range = 24-60).

Tutee 5

Tutee 5 exhibited the most variable data set across all tutees. His fluctuation around mean score during baseline showed the highest levels of variability among tutees (SD = 7.26; range = 17-35; See Table 4.1). During intervention his standard deviation stabilized, compared to baseline (SD = 3.85; range = 27-39). Tutee 5's TEMI-AC total scores during baseline demonstrated an overall upward trend (0.24; See Figure 4.4, Panel 5). Although his baseline trend was not stable, a decision was made to introduce the intervention due to time constraints. Tutee 5 continued to show an upward trend (0.46) during intervention, but at a rate almost twice that of baseline's trend. His total score level during baseline was 25.7 and increased during intervention to 32.2 (See Figure 4.3, Panel 5). Tutee 5 was the only tutee to demonstrate a negative immediacy of effect (-13.33%) and was also one of two tutees to have overlapping data across phases. Maintenance phase showed an increase in level (42.5) compared to intervention and baseline phases, and both the 2-week and 4-week maintenance data points exceeded all TEMI-AC total scores within the baseline and intervention phases.

Summary

In sum, once the intervention was introduced, three of five tutees showed an immediacy of effect, at high rates, from baseline to intervention phase (range = 73.33 - 116.67%). Two of five tutees were slower to respond to the intervention and showed overlap data across phases. This overlap in data frequently occurred within the first half of the intervention phase. The average change in level from baseline to intervention was 55.33% across tutees. Trend analysis showed three tutees with stable baselines prior to the intervention and then sequential upward trends during the intervention phase (range = 0.55-1.0). For the two tutees showing upward trends during baseline (0.24; 0.34), once the intervention was introduced they also demonstrated upward trends in scores at almost double the rate, in both cases (range = 0.46-0.89). The mean variability (i.e., standard deviation) across tutee TEMI-AC weekly total scores was 4.21 (SD = 2.04) during baseline and 5.27 (SD = 3.42) during intervention. All tutees that were available to be assessed for maintenance of mathematics performance showed increases in level from

Tutees	Intervention[T]	Maintenance[T]

that of baseline and intervention phases, showing a retention of skills and knowledge as well as continuing development after the intervention was removed. Based on the visual analysis findings, a causal relation was demonstrated between the cross-age tutoring program containing number line board games and the mathematics performance of at-risk kindergartens on early numeracy measures.

Effect Sizes: Proximal Data (TEMI-AC)

In addition to visual analyzing the data, the percentage of data points exceeding the median of baseline phase (PEM; Ma, 2006) was calculated to assess the magnitude and strength of effects. PEM results are analyzed using the following scale: 90-100% = large or highly effective, 70%-90% = moderately effective, and < 70% = small or questionable effectiveness (Ma, 2006). Table 4.2 shows PEM effect sizes for intervention and maintenance phases.

Tutee 1	90%	100%
Tutee 2	100%	100%
Tutee 3	100%	N/A
Tutee 4	100%	100%
Tutee 5	80%	100%
Mean (SD)	94% (8.9)	100% (0.0)

Note. [T] = TEMI-AC total score; * = completed only 3 weeks of intervention

Table 4.2: Tutees' Percentage of Data Points Overlapping Median (PEM) by Phase.

Tutee 1

From baseline to intervention phase, Tutee 1's data demonstrated a PEM effect size of 90% on TEMI-AC total scores. There was 10% of data points within the intervention phase that overlapped with the median data point within the baseline phase, and 0% overlap within the maintenance phase (PEM = 100%). The effect of the tutoring intervention on TEMI-AC scores can be interpreted as showing moderate effectiveness during the intervention phase and having large effect during the maintenance phase, compared to baseline (Ma, 2006).

Tutee 2

According to PEM (100%) data, Tutee 2 improved his TEMI-AC scores 100% between baseline and intervention phase; there was no overlap data between phases. The PEM value demonstrated a large effect of the tutoring program on Tutee 2's performance on weekly TEMI-AC probes. During maintenance phase, PEM (100%) demonstrated lasting large effects on TEMI-AC scores at 2- and 4-weeks after the last tutoring session was attended.

Tutee 3

Tutee 3 only completed three weeks of the intervention phase and all three TEMI-AC total scores within this phase were above the median baseline point (PEM = 100%), demonstrating large effect of the tutoring intervention within the shortened duration. Maintenance data was unavailable for this tutee due to their exit from the program after the third week.

Tutee 4

According to Tutee 4's PEM data (100%), he showed 100% improvement in TEMI-AC weekly scores compared to the median baseline score. This is interpreted to be a large effect size for the intervention phase. During maintenance, Tutee 4 continued to score well above the median baseline level and continued to demonstrate the tutoring intervention's large effect on the tutee's performance on TEMI-AC probes at 2- and 4-weeks after the intervention was removed.

Tutee 5

Tutee 5's PEM data demonstrated an 80% improvement in TEMI-AC total scores from baseline to intervention, with 20% of data points overlapping the median baseline score. This PEM data is interpreted to show moderate effectiveness of the tutoring sessions on Tutee 5's weekly mathematics performance on TEMI-Ac probes.
Maintenance PEM data (100%) showed the intervention's large effect on TEMI-AC scores after the last tutoring session, compared to baseline.

Overall

Overall PEM data showed moderate to large effects of the tutoring intervention on tutees' weekly mathematics performance on TEMI–AC probes during the intervention phase. Across tutees, the PEM effect size mean was 94% (SD = 8.9), which is interpreted to demonstrate a highly effective intervention. After the final tutoring session of the program, all tutees demonstrated better performance on TEMI-AC probes during maintenance phase compared to median baseline scores (PEM = 100%; SD = 0.0), exhibiting the intervention's large maintaining effects.

Effect Sizes: Pre-/Post-intervention Measures (TEMA-3)

In addition to calculating and evaluating proximal effects of the intervention, a distal measure, the TEMA-3 was also administered. Pre-intervention TEMA-3 scores were obtained during screening procedures. To assist in assessing generalized outcomes, Hedges' *g* effect size was calculated. Effect sizes are interpreted based on the criteria set by Cohen (1969), where 0.2 - 0.49 is interpreted as a small effect, 0.5 - 0.79 shows medium or moderate effect, and 0.8 or greater suggests large or significant effect. Table 4.3 shows pre-/post-intervention, distal measure data for tutees and tutors.

Tutees	TEMA-3 *Pre*	TEMA-3 *Post*
Tutee 1	10	26
Tutee 2	8	17
Tutee 3	11	N/A
Tutee 4	9	21
Tutee 5	6	15
Mean (SD)	8.8 (1.9)	19.8 (4.9)
Effect size (*g*)	*g* = 2.78	

Note. TEMA–3 = Test of Early Mathematics Ability–3: Raw score (Ginsburg & Baroody, 2003); SD = standard deviation

Table 4.3: Tutees' Pre-/Post-intervention Scores and Effect Size.

TEMA-3 was administered prior to the start of the intervention and tutees' raw scores were low, ranging from 6 to 11 (SD = 1.9). These scores fell within the 'below average' to 'poor' TEMA-3 categorizations for mathematics ability levels. The mean pre-test score across tutees was 8.8, which falls within the 'poor' category (i.e., requiring intervention for mathematics difficulties). After ten weeks of intervention, the students entered maintenance phase (i.e., no longer attended tutoring sessions) and the TEMA-3 was administered as a post-test. The mean raw score across tutees was found to be 19.8 (SD = 4.9), with scores ranging from 15 to 26. The cross-age tutoring program was found to have a large, statistically significant effect (g = 2.78) on tutees' mathematics performance from pre- to post-intervention.

Summary

Research question one, which assessed the extent to which the cross-age tutoring intervention was effective in promoting the mathematics performance of kindergarteners at-risk for disability, was evaluated through three sets of analysis: visual analysis, proximal effect sizes (i.e., of visual data), and distal effect sizes (i.e., pre-/post-assessment). First, visual analysis showed that a causal relation was demonstrated between the cross-age tutoring program containing number line board games and the mathematics performance of at-risk kindergartens on early numeracy measures. Additionally, across tutees, performance on TEMI-AC probes showed maintenance of scores at 2- and 4-weeks after the intervention. Second, evaluation of effect sizes of visual data (i.e., PEM) showed moderate to large effects, with the mean effect size demonstrating large improvements across tutees' performance on weekly probes. Furthermore, all tutees demonstrated improved performance during maintenance phase compared to baseline scores, which also yielded a large effect size. Third, distal data was examined and demonstrated statistically significant effects (g = 2.78) on tutees' mathematics performance on TEMA-3 from pre- to post-intervention. In sum, through multiple analyses, the cross-age tutoring program demonstrated effectiveness in promoting at-risk kindergarten students' mathematics performance on both proximal and distal measures of early numeracy knowledge and skills.

RESEARCH QUESTION 2

Research question two assessed the ability of students with EBD to effectively serve as cross-age tutors. This was evaluated through two components, the first being the tutors' ability implement instructional procedures with fidelity during the intervention phase. The second evaluating component was the tutors' effectiveness in promoting tutee performance (i.e., increasing scores) on proximal and distal measures of early numeracy knowledge and skills.

The first component (i.e., fidelity of implementation) was assessed through the completion of the fidelity checklist during observations of the cross-age tutors during each tutor session. The checklist measured the tutors' level of implementation of the tutoring procedures that they were trained upon prior to the start of the intervention. Table 4.4 shows tutors' fidelity of implementation percentages of tutoring procedures throughout the intervention phase as well as means and standard deviations across tutors and across weeks.

Week	1	2	3	4	5	6	7	8	9	10	**Overall mean (SD)**
Tutor 1	69*	87.5	79.3*	96	100	97.3	100	91.7	100	97.3	91.8 (10.4)
Tutor 2	87.7	91.5	91.7	96	100	100	100	96	94.3	97.3	95.5 (4.2)
Tutor 3	80	79.3*	91.5	N/A	N/A	N/A	N/A	N/A	N/A	N/A	83.6 (6.9)
Tutor 4	69*	80	91.7	96	96	97.3	100	91.7	90.3	97.3	90.9 (9.5)
Tutor 5	66.3*	80.8	91.7	94.7	97.3	94.3	92	96	100	93.3	90.6 (9.9)
Weekly mean (SD)	74.4 (9.1)	83.8 (5.4)	89.2 (5.5)	95.7 (0.7)	98.3 (2.0)	97.2 (2.3)	98 (4.0)	93.9 (2.5)	96.2 (4.7)	96.3 (2.0)	**90.5 (4.32)**

Note. SD = standard deviation; * = retraining session provided (i.e., weekly *M* < 80%)

Table 4.4: Tutors' Fidelity of Implementation Percentages during Intervention.

Tutor 1's fidelity of implementation ranged from 69% to 100% throughout the intervention and averaged at a rate of 91.8% (SD = 10.4). He required a retraining session (i.e., due to the weekly fidelity mean falling below the criteria of 80%) during intervention week one and week three. After each of the retraining sessions, his implementation rate in the following weeks to 87.5% and 96%, respectively. After the first three weeks of implementation, Tutor 1 did not require any further retraining sessions and showed an upward trend in fidelity of implementation rates for the remaining weeks of the intervention.

Tutor 2 had an overall fidelity average of 95.5% (SD = 4.2) during the intervention (Range = 87.7% to 100%). Tutor 2 did not require any additional retraining sessions, as his rate of implementation was stable above 80% throughout the 10 weeks of intervention. His average level of implementation was above 90% for nine out of ten weeks of the intervention.

Tutor 3, who was only able to implement the tutoring procedures for three weeks due to his tutee moving out of district, had an average fidelity rate of 83.6%, with levels ranging from 79.3% to 91.5% (SD = 6.9). Tutor 3 required one retraining session due to a weekly average during the second week that fell below the 80% criteria. After the retraining session, the tutor's fidelity of implementation increased to 91.5%.

Tutor 4's fidelity of implementation ranged from 69% to 100%, with a mean of 90.9% (SD = 9.5) throughout the intervention phase. One retraining session was required for Tutor 4, due to his weekly fidelity rate of 69% during week one. After the retraining session was provided the tutor was able to implement the tutoring instructional procedures at a fidelity rate greater than 80% for the remaining weeks of the intervention

Tutor 5 had an overall mean fidelity rate of 90.6% (SD = 9.9), with weekly means ranging from 66.3% to 100% across the ten weeks of intervention. Tutor 5 also required one retraining session after the first week of intervention in which his weekly fidelity average was 66.3%. After retraining his rate increased to 80.8% in week two, 91.7% in week three, and subsequently had an overall increasing trend throughout the remaining weeks of the intervention.

Overall, across tutors, the mean fidelity of implementation during the intervention was 90.5% (SD = 4.32). Weekly means across tutors were above 90% in seven of ten intervention weeks, and above 80% in nine out of ten intervention weeks. Across tutors, there was a steady increasing trend in the weekly fidelity means from week one onward through week five and then stabilized above 95% for all of the remaining weeks of the intervention. Retraining sessions were required for four of five tutors, and all retraining sessions took place within the first three weeks of implementation. No tutor required a retraining session after the third week of intervention.

The second component (i.e., tutee outcomes on early numeracy measures) was addressed through research question one, which assessed the effectiveness of the intervention on promoting tutees' performance on mathematics measures. Results related to research question one showed moderate to large effects on weekly mathematics scores on TEMI–AC probes from baseline to intervention phases (PEM mean = 94%; range = 80%–100%). Furthermore, on distal measures, tutees' showed statically significant gains on pre-/post-intervention TEMA–3 scores (g = 2.78). The positive gains in tutees' mathematical performance on proximal and distal measures show that the tutors' were effective in promoting the intervention's targeted skills and knowledge for the tutees.

Summary

The ability to implement tutoring instructional procedures with fidelity during intervention, and also the ability to increase early numeracy knowledge and skills of tutees were used as evaluating components to addressing the question of whether students with EBD can effectively serve as cross-age tutors. In assessing these two components, results showed that tutors were able to both implement tutoring instructional procedures that they were trained upon, with high rates of fidelity throughout the intervention (across tutors: M = 90.5%; range = 83.6%–95.5%), as well as large and statistically significant gains in mathematical performance of tutees as assessed by performance on early numeracy measures. The combination of these two results demonstrates the ability of students with/at-risk for EBD to effectively serve as cross-age tutors.

RESEARCH QUESTION 3

Research question 3 examined the effects of a number line board game, delivered by a cross-age tutor with EBD, on the mathematics performance of kindergarten students at-risk for math disabilities. To assess the early numeracy knowledge and skills of these students at-risk for mathematics disabilities (i.e., tutees) weekly probes were administered in the form of Texas Early Mathematics Inventory–Aim Checks (TEMI-AC; University of Texas System/Texas Education Agency, 2009). The four subtest scores were combined to create a total score, which was then used to evaluate students' progress on a weekly basis.

Visual Analysis

A visual analysis was conducted according to the recommendations for evaluating the essential features of single-case design data (Kratochwill et al., 2010) to determine if a casual relation existed between the tutor training and the attendance and implementation of the cross-age tutoring program (Independent Variable) and the tutors' behavioral ratings as measured through their daily Check-in/Check-out (CICO) behavior point sheet scores, averaged by week (Dependent Variable). Figure 4.5 displays tutors' weekly average for CICO behavior points obtained. Figure 4.6 shows level data and Figure 4.7 shows trend data for tutors' CICO weekly mean scores, along with weekly average scores disaggregated for tutoring and non-tutoring days. Table 4.5 shows variability, immediacy of effect, and overlap data for tutors.

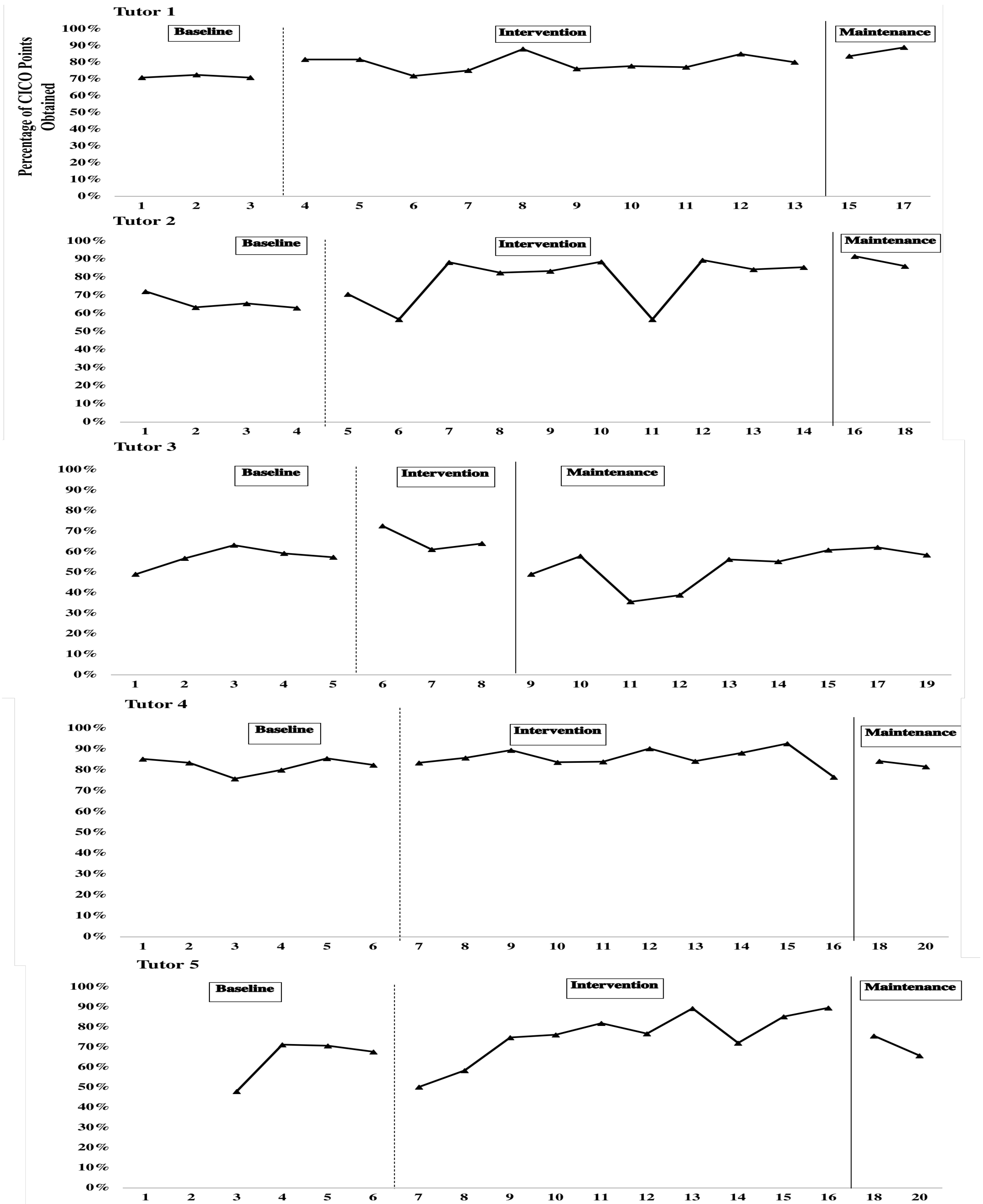
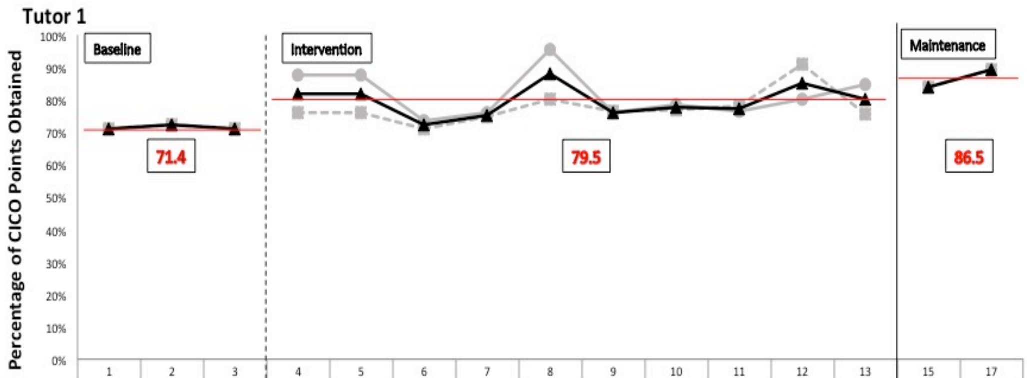


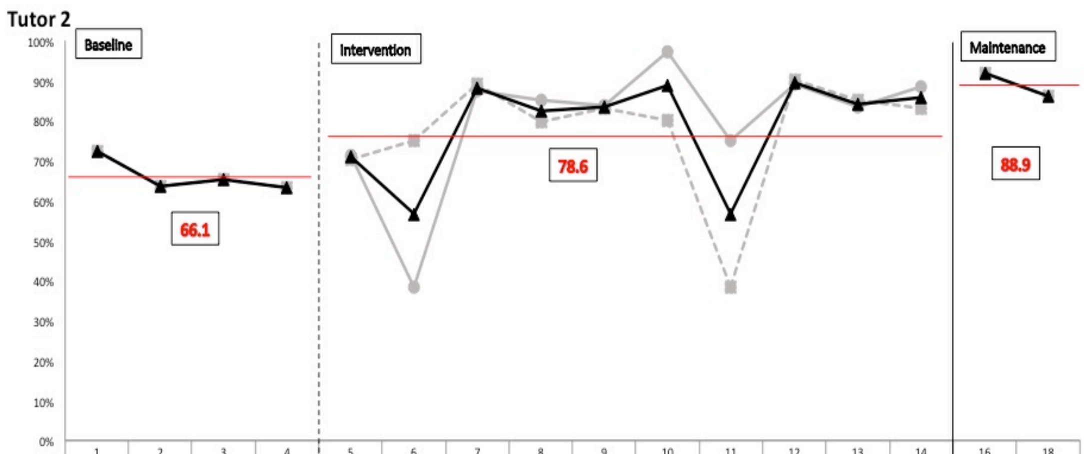
Figure 4.5: Tutors' Average CICO Behavioral Points Obtained by Week

Tutor 1



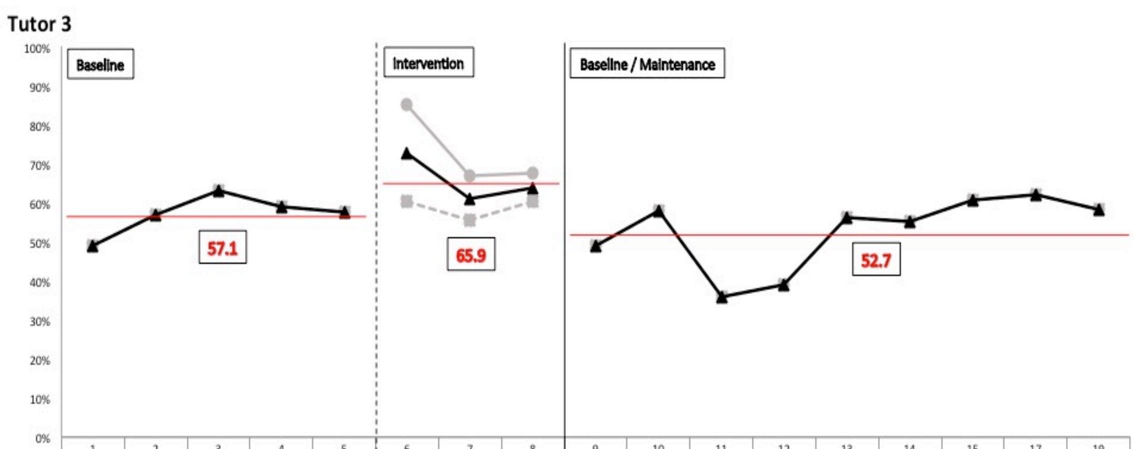
	1	2	3	4	5	6	7	8	9	10	11	12	13	15	17
TUTORING DAYS-Weekly Mean				87.5%	87.7%	73.3%	75.8%	95.6%	75.9%	78.7%	76.3%	80.0%	84.8%		
NON-TUTORING DAYS - Weekly Mean	70.8%	72.5%	71.0%	75.8%	75.8%	70.8%	74.6%	80.3%	76.3%	76.9%	78.2%	91.0%	75.6%	83.7%	89.2%
OVERALL - Weekly Mean	70.8%	72.5%	71.0%	81.7%	81.6%	72.1%	75.2%	88.0%	76.1%	77.8%	77.3%	85.1%	80.2%	83.7%	89.2%

Tutor 2

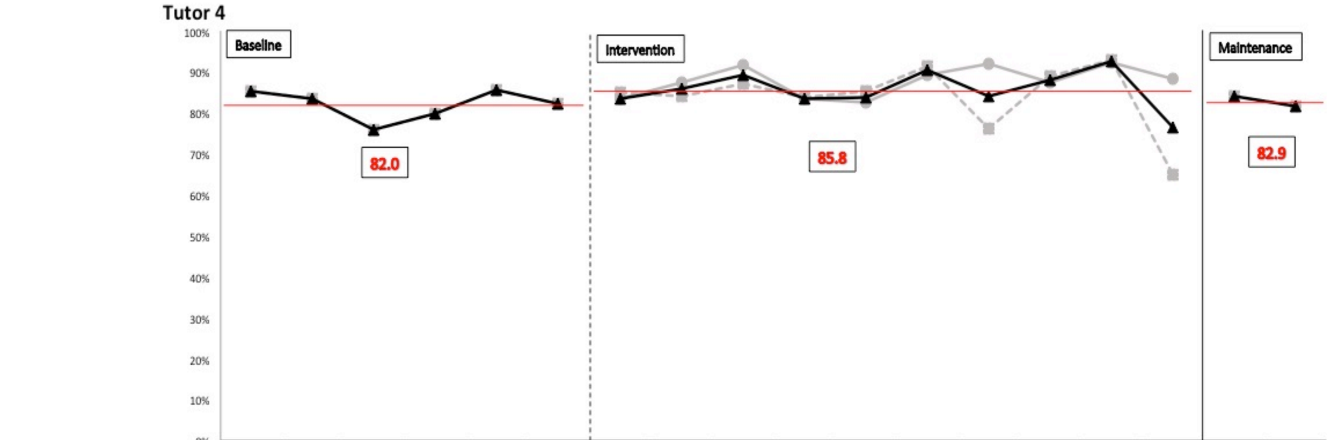


	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	18
TUTORING DAYS-Weekly Mean					71.2%	38.3%	87.3%	85.0%	83.7%	97.2%	75.0%	89.1%	83.4%	88.3%		
NON-TUTORING DAYS - Weekly Mean	72.3%	63.4%	65.4%	63.1%	70.4%	75.0%	89.1%	79.8%	83.2%	80.0%	38.3%	90.0%	85.0%	82.9%	91.7%	86.1%
OVERALL - Weekly Mean	72.3%	63.4%	65.4%	63.1%	70.8%	56.7%	88.2%	82.4%	83.5%	88.6%	56.7%	89.6%	84.2%	85.6%	91.7%	86.1%

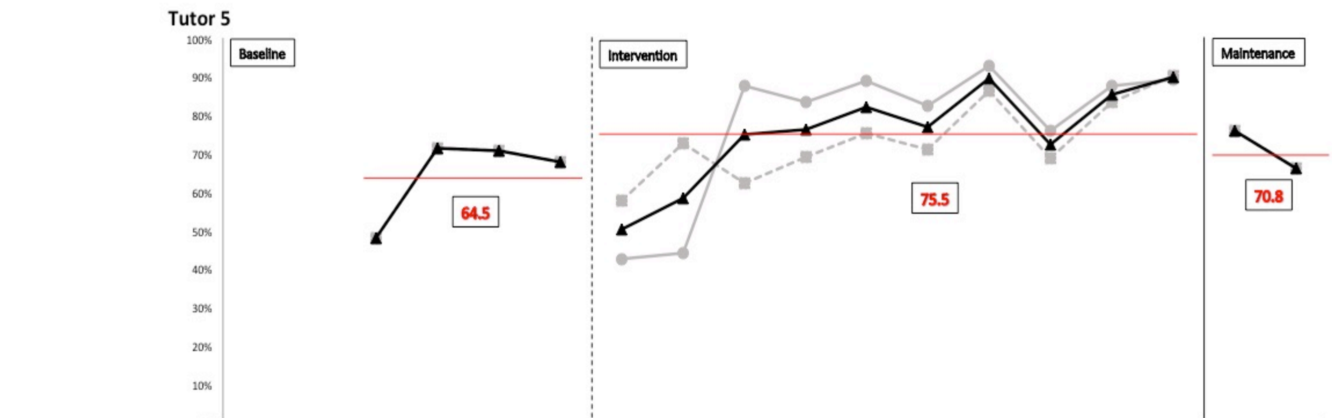
Tutor 3



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	17	19
TUTORING DAYS-Weekly Mean						85.1%	66.7%	67.4%									
NON-TUTORING DAYS - Weekly Mean	49.0%	56.7%	63.2%	59.1%	57.4%	60.3%	55.6%	60.4%	49.1%	57.9%	35.7%	38.9%	56.3%	55.3%	60.8%	62.1%	58.3%
OVERALL - Weekly Mean	49.0%	56.7%	63.2%	59.1%	57.4%	72.7%	61.2%	63.9%	49.1%	57.9%	35.7%	38.9%	56.3%	55.3%	60.8%	62.1%	58.3%



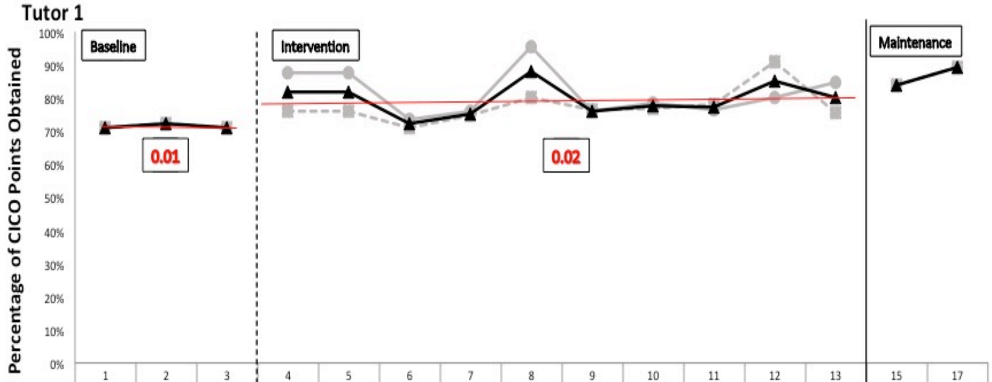
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
TUTORING DAYS-Weekly Mean							83.3%	87.3%	91.7%	83.3%	82.4%	89.1%	92.1%	87.3%	92.4%	88.2%		
NON-TUTORING DAYS - Weekly Mean	85.2%	83.3%	75.9%	79.9%	85.6%	82.3%	84.9%	84.2%	87.0%	83.8%	85.3%	91.5%	76.2%	89.0%	93.0%	65.0%	84.2%	81.5%
OVERALL - Weekly Mean	85.2%	83.3%	75.9%	79.9%	85.6%	82.3%	83.4%	85.8%	89.4%	83.6%	83.9%	90.3%	84.2%	88.2%	92.7%	76.6%	84.2%	81.5%



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
TUTORING DAYS-Weekly Mean							42.5%	44.2%	87.4%	83.3%	88.7%	82.4%	92.8%	75.8%	87.6%	89.2%		
NON-TUTORING DAYS - Weekly Mean			48.0%	71.4%	70.8%	67.8%	57.8%	72.5%	62.4%	69.2%	75.3%	71.1%	86.2%	68.6%	83.2%	90.0%	75.7%	65.9%
OVERALL - Weekly Mean			48.0%	71.4%	70.8%	67.8%	50.1%	58.3%	74.9%	76.3%	82.0%	76.8%	89.5%	72.2%	85.4%	89.7%	75.7%	65.9%

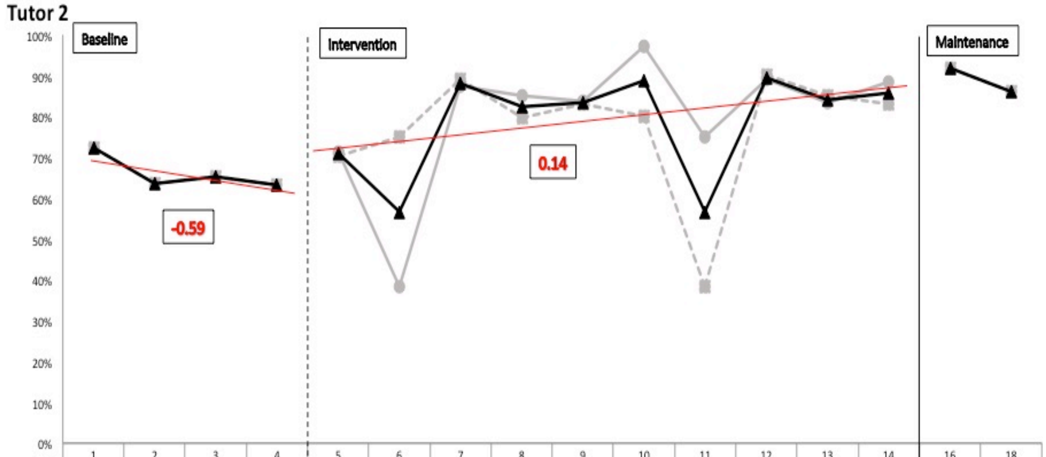
Figure 4.6: Level and Disaggregated Data for Tutors' Weekly CICO Scores

Tutor 1



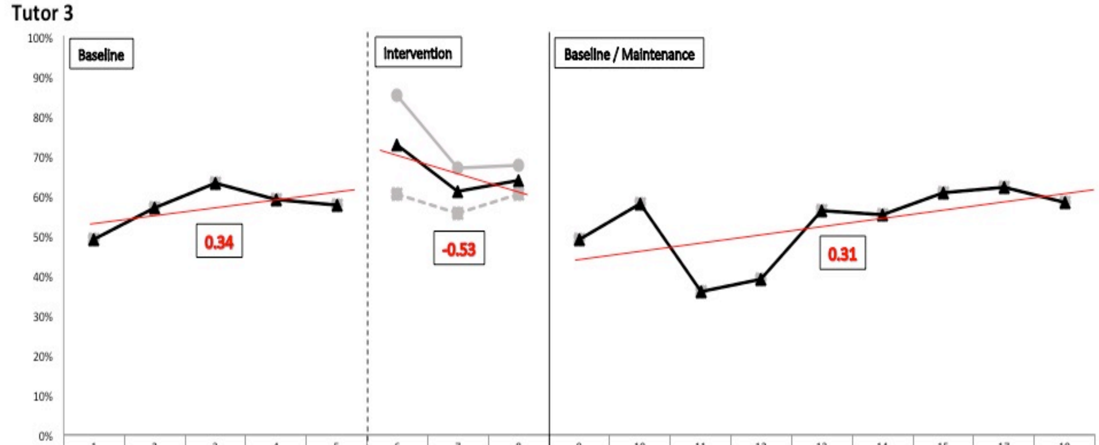
	1	2	3	4	5	6	7	8	9	10	11	12	13	15	17
TUTORING DAYS-Weekly Mean				87.5%	87.7%	73.3%	75.8%	95.6%	75.9%	78.7%	76.3%	80.0%	84.8%		
NON-TUTORING DAYS - Weekly Mean	70.8%	72.5%	71.0%	75.8%	75.8%	70.8%	74.6%	80.3%	76.3%	76.9%	78.2%	91.0%	75.6%	83.7%	89.2%
OVERALL - Weekly Mean	70.8%	72.5%	71.0%	81.7%	81.6%	72.1%	75.2%	88.0%	76.1%	77.8%	77.3%	85.1%	80.2%	83.7%	89.2%

Tutor 2



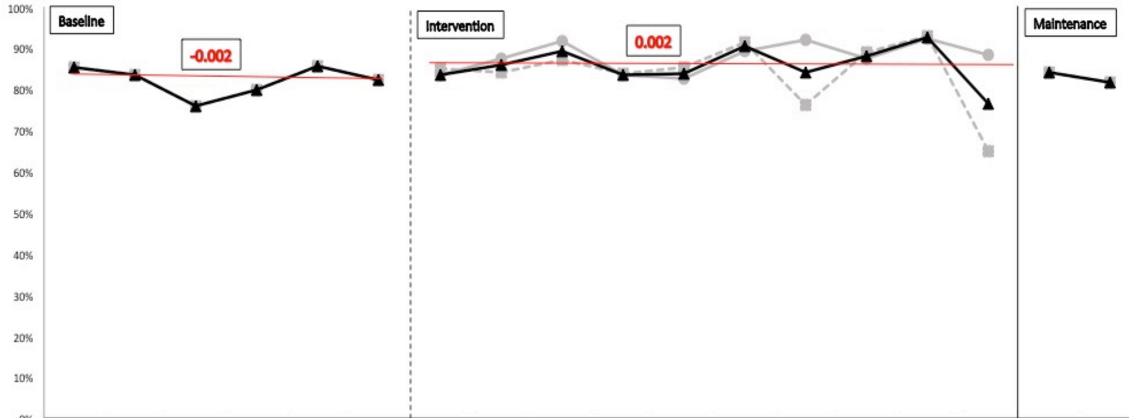
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	18
TUTORING DAYS-Weekly Mean					71.2%	38.3%	87.3%	85.0%	83.7%	97.2%	75.0%	89.1%	83.4%	88.3%		
NON-TUTORING DAYS - Weekly Mean	72.3%	63.4%	65.4%	63.1%	70.4%	75.0%	89.1%	79.8%	83.2%	80.0%	38.3%	90.0%	85.0%	82.9%	91.7%	86.1%
OVERALL - Weekly Mean	72.3%	63.4%	65.4%	63.1%	70.8%	56.7%	88.2%	82.4%	83.5%	88.6%	56.7%	89.6%	84.2%	85.6%	91.7%	86.1%

Tutor 3



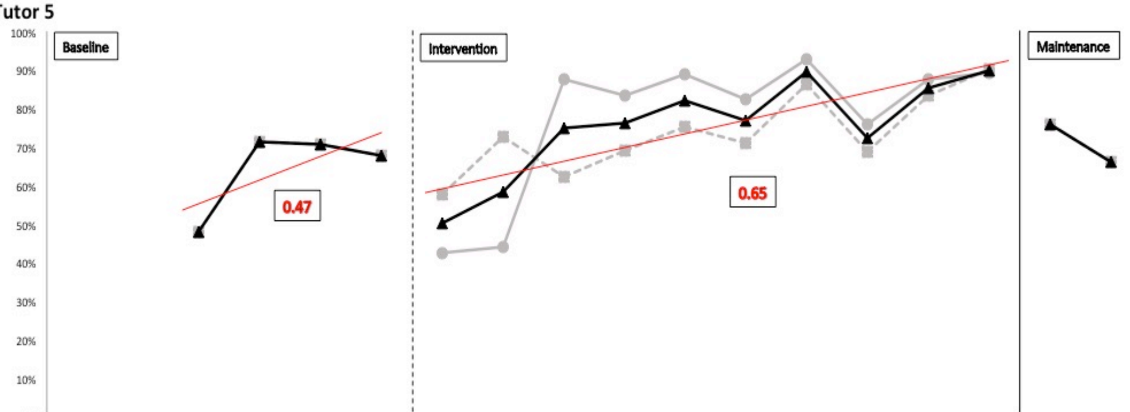
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	17	19
TUTORING DAYS-Weekly Mean						85.1%	66.7%	67.4%									
NON-TUTORING DAYS - Weekly Mean	49.0%	56.7%	63.2%	59.1%	57.4%	60.3%	55.6%	60.4%	49.1%	57.9%	35.7%	38.9%	56.3%	55.3%	60.8%	62.1%	58.3%
OVERALL - Weekly Mean	49.0%	56.7%	63.2%	59.1%	57.4%	72.7%	61.2%	63.9%	49.1%	57.9%	35.7%	38.9%	56.3%	55.3%	60.8%	62.1%	58.3%

Tutor 4



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
TUTORING DAYS-Weekly Mean							83.3%	87.3%	91.7%	83.3%	82.4%	89.1%	92.1%	87.3%	92.4%	88.2%		
NON-TUTORING DAYS - Weekly Mean	85.2%	83.3%	75.9%	79.9%	85.6%	82.3%	84.9%	84.2%	87.0%	83.8%	85.3%	91.5%	76.2%	89.0%	93.0%	65.0%	84.2%	81.5%
OVERALL - Weekly Mean	85.2%	83.3%	75.9%	79.9%	85.6%	82.3%	83.4%	85.8%	89.4%	83.6%	83.9%	90.3%	84.2%	88.2%	92.7%	76.6%	84.2%	81.5%

Tutor 5



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
TUTORING DAYS-Weekly Mean							42.5%	44.2%	87.4%	83.3%	88.7%	82.4%	92.8%	75.8%	87.6%	89.2%		
NON-TUTORING DAYS - Weekly Mean			48.0%	71.4%	70.8%	67.8%	57.8%	72.5%	62.4%	69.2%	75.3%	71.1%	86.2%	68.6%	83.2%	90.0%	75.7%	65.9%
OVERALL - Weekly Mean			48.0%	71.4%	70.8%	67.8%	50.1%	58.3%	74.9%	76.3%	82.0%	76.8%	89.5%	72.2%	85.4%	89.7%	75.7%	65.9%

Figure 4.7: Trend and Disaggregated Data for Tutors’ Weekly CICO Scores

	Baseline°	Intervention°		
Tutor 1	0.90 (70.8-72.5)	4.78 (75.2-88.0)	71.0	No
Tutor 2	4.29 (63.1-72.3)	1.27 (56.7-89.6)	79.0	Yes
Tutor 3	5.17 (49.0-63.2)	6.03 (35.7-62.1)	60.0	No
Tutor 4	3.65 (75.9-85.6)	4.57 (76.6-92.7)	36.0	Yes
Tutor 5	11.11 (48.0-71.4)	12.86 (50.1-89.7)	-89.0	Yes
Mean SD (SD)	5.02 (3.76)	5.90 (4.27)	31.4 (69.23)	

Note. ° = overall weekly CICO mean; SD = standard deviation

Table 4.5: Tutors' Variability, Immediacy of Effect, and Overlap data.

Tutor 1

Tutor 1 had a stable baseline trend (0.01) and his CICO weekly means increased in level from baseline to intervention phase from 71.4 to 79.5 (see Figure 4.6, Panel 1), and had the second highest immediacy of effect (71%) across tutors. Overlap data was not present across phases. The trend during intervention phase was also, relatively stable (0.02; see Figure 4.7, Panel 1) but increased in variability (SD = 4.78; range = 75.2-88.0; see Table 4.5) compared to baseline data, which was the most stable across tutors (SD = 0.90; range = 70.8-72.5). Maintenance phase showed a greater level (86.5) than both baseline and intervention phases, as well as a slight increasing trend.

Tutor 2

Tutor 2 had the highest immediacy of effect (79%) among all tutors. His level during baseline was 66.1 (see Figure 4.6, Panel 2) with a small decreasing trend (-0.59;

see Figure 4.7, Panel 2). During intervention, the level increased to 78.6 and demonstrated an increasing trend (0.14). Tutor 2 also had the most significant increase in level from intervention to maintenance and ended the phase with a level of 88.9, which was the highest across tutors. Fluctuation around mean score was relatively average compared to other tutors (SD = 4.29; range = 63.1-72.3) and stabilized further during intervention (SD = 1.27; range = 56.7-89.6), although there were two low, outlier CICO scores, which may account for the overlap in data across phases.

Tutor 3

Due to Tutee 3 moving out of district after the third intervention week, tutoring program responsibilities and attendance at sessions were removed for Tutor 3, and his CICO data returned to baseline phase. This phase can also be interpreted as an extended maintenance phase, lasting for nine consecutive weeks. As previously stated, this decision was made for two reasons: first, to assess any change in data patterns and, second, to determine any potential long-term maintenance effects on CICO behavioral scores.

His initial level during baseline was 57.1, increased to 65.9 during intervention, and then decreased to 52.7 when the intervention was removed (i.e., baseline/maintenance; see Figure 4.6, Panel 3). During baseline, data exhibited a slight upward trend (0.34), but was stable across the final four data points within the phase. Once the intervention was introduced, the data showed a downward trend (-0.53), with the highest CICO mean coming during the first week of the intervention, although it considerations need to be made for the limited number of data points within this phase. Additionally, in evaluating the three CICO mean scores during intervention, it is noted that they are all at or above the highest baseline CICO mean (see Figure 4.7, Panel 3).

Tutor 3 exhibited the second greatest immediacy of effect across all tutors (60%), with no overlapping data. Baseline variability was found to be average compared to other tutees (SD = 5.17; range = 49.0-63.2), while intervention variability was slightly higher (SD = 6.03; range = 35.7-62.1). Maintenance phase (i.e., reverted to baseline phase after week 3) lasted nine consecutive weeks and had an overall level of 52.7, which decreased immediately after the intervention was removed. The data within this extended maintenance phase showed an upward trend (0.31) and also contained a high level of variability in CICO means across weeks.

Tutor 4

Tutor 4's CICO performance showed the highest baseline level (82.0; see Figure 4.6, Panel 4) and also the most stable phase (-0.002; see Figure 4.7, Panel 4) across tutors. His variability in performance during baseline was relatively average (SD = 3.65; range = 75.9-85.6) and remained so during intervention (SD = 4.57; range = 76.6-92.7). Tutor 4's CICO weekly means demonstrated an immediacy of effect upon the introduction of the intervention (36%) but still contained overlapping data. During the intervention phase, his data was stable, with a slight upward trend (0.002). Maintenance level was 82.9, which was above his baseline level but slightly below his intervention phase level.

Tutor 5

Most of the tutors showed stable baseline trends with the exception of Tutor 5 who showed an upward trend (0.47), although accounting for the first outlier score, the three scores immediately prior to the introduction of intervention showed a stable trend line (see Figure 4.7, Panel 5). Tutor 5 was also the only tutor to have an initial change in level that was negative (-89%), although his overall level during intervention (75.5) was higher than his baseline level (64.5). Tutor 5 had a baseline standard deviation of 11.11 (range = 48.0-71.4), which was the most variable across all tutors. This trend continued into the intervention phase, where his standard deviation was SD = 12.86 (range = 50.1-89.7). Overlap of data was present but only during the first two weeks of the intervention phase. From week three onward, his CICO weekly means exceeded baseline and continued with an upward trend during the following weeks (0.65). Finally, Tutor 5 also was one of two tutors to have a CICO weekly mean score maintenance level (70.8) that was below the intervention level (75.5), but still higher than the level during the baseline phase (64.5).

Summary

In sum, all five tutors' data demonstrated a positive change in level from baseline to intervention phase. Immediacy of effect was present in 4/5 tutors; mean 31.4% increase change in level in CICO scores (range = -89%-79%). From intervention to maintenance phase, two tutors showed an increase in level, while two tutors showed slight decreases in level but remained higher than initial baseline levels in both cases. Tutor 3, who only attended three weeks of intervention (i.e., tutee moved out of district), returned to roughly the same level as baseline phase when the intervention was removed.

In assessing the overall trend in data, and accounting for Tutor 5's outlier baseline score, all tutors showed stable baselines, and once the intervention was introduced, three of five tutors showed upward trends, one tutor (i.e., Tutor 4) showed no directional trend, and one tutor's data demonstrated a downward trend (-0.53). It should be noted that this was the tutor (i.e., Tutor 3) who only participated in three tutoring sessions before returning to baseline phase, and additionally, all three of his intervention data points were at, or above his highest baseline data point. Furthermore, once the intervention was removed, the tutors' CICO scores decreased in level and remained that of the intervention phase mean.

Three of five tutors showed overlap of data from baseline to intervention phase. This variability can be seen in the baseline and intervention standard deviations across tutors' weekly CICO mean scores, which ranged from 0.90 to 11.11 (*M* = 5.02; SD = 3.76) and 4.57 to 12.86 (*M* = 5.9; SD = 4.27), respectively. Maintenance phases across tutors showed all tutors to maintain or increase behavioral score levels at, or above intervention levels, and all maintenance data points, across tutors, were well above that of baseline levels. Based on the visual analysis of the data, a causal relation could not be fully demonstrated between the cross-age tutoring program training and implementation, and the tutors' behavioral performance as measured by CICO point sheets due to the high rates in variability in data of two of the five participants.

Effect Sizes: Proximal Data (CICO)

Percentage of data points exceeding the median of the baseline phase (PEM; Ma, 2006) was also calculated to assess the magnitude and strength of effects for tutors' performance on CICO point sheets from baseline through intervention phases. PEM results are analyzed using the following scale: 90-100% = large or highly effective, 70%-90% = moderately effective, and < 70% = small or questionable effectiveness (Ma, 2006).

Note. T = tutoring days; NT = non-tutoring days; O = overall weekly mean; * = returned to baseline after intervention week 3

Table 4.6: Tutors' Percentage of Data Points Overlapping Median (PEM) by Phase.

Tutors	Intervention			Maintenance
	T	NT	O	
Tutor 1	100%	90%	100%	100%
Tutor 2	90%	90%	80%	100%
Tutor 3	100%	67%	100%	44%*
Tutor 4	90%	80%	90%	50%
Tutor 5	80%	60%	80%	50%
Mean (SD)	92% (8.4)	77.4% (13.6)	90% (10.0)	68.8% (28.6)

Tutor 1

From baseline to intervention phase, Tutor 1's PEM data demonstrated an improvement of 100% on days when he attended tutoring sessions (i.e., served as a tutor), 90% on non-tutoring days, and an overall weekly mean of 100% (i.e., mean of tutoring and non-tutoring day CICO scores). There was an overlap of 10% of CICO scores on non-tutoring days from baseline to intervention phase. The effect of tutor training and implementation of tutoring sessions on Tutor 1's CICO behavioral scores is interpreted to be large on tutoring days, moderate on non-tutoring days, and large on overall (i.e., tutoring and non-tutoring days) weekly mean scores (Ma, 2006). During the maintenance phase, Tutor 1's scores remained at levels consistent during tutoring days and well above baseline weekly mean scores.

Tutor 2

According to Tutor 2's PEM data, he improved his performance on CICO behavioral point sheets by 90% on tutoring days and 90% on non-tutoring days. Overall PEM, on weekly mean CICO scores, showed 80% improvement in weekly behavioral ratings across days. These results demonstrate the tutoring program's moderate effect on improving tutors' behavior, as measured by CICO scores. During maintenance phase, PEM (100%) demonstrated lasting large effects on weekly CICO point averages 2- and 4-weeks after the last tutoring session was attended.

Tutor 3

Tutor 3's tutee left the district after week three of the intervention and he was forced to return to baseline phase, which can also be evaluated as an extended maintenance phase of nine consecutive weeks. From baseline to intervention phase, PEM data showed an improvement in CICO scores of 100% on tutoring days and 67% on non-tutoring days. Tutor 3's non-tutoring day data during intervention demonstrated the highest rates of overlap data compared to other tutors, with 33% of weekly CICO scores falling at or below the median score during baseline. Overall weekly mean scores during intervention exhibited PEM of 100% in relation to baseline. Tutor 3's CICO scores decreased immediately following the removal the intervention and stay consistently low when he returned to baseline. Evaluating the unique maintenance PEM data of Tutor 3 (i.e., he had seven additional data points compared to other tutors) found the lowest effect size of any tutor during this phase, with just 44% of CICO scores being greater than the baseline median score. This can be interpreted in that the tutoring program showed small, or questionable effects after the intervention was prematurely removed due to unforeseen circumstances.

Tutor 4

According to Tutor 4's PEM data, he showed 90% improvement in CICO scores during tutoring days, 80% improvement on non-tutoring days, and 90% improvement on overall weekly average scores, compared to the median baseline score. This is interpreted to show the tutoring program's moderate effects on CICO score performance across days of the intervention phase. Once the intervention was removed and the maintenance phase began, Tutor 4's CICO weekly scores decreased and exhibited 50% overlap in data with his median baseline score. This showed the intervention's questionable effectiveness in providing lasting effects on Tutor 4's behavioral ratings when he was not attending tutoring sessions during the week.

Tutor 5

Tutee 5's PEM data demonstrated an 80% improvement in CICO scores on tutoring days, 60% improvement on non-tutoring days and 80% improvement on overall weekly mean scores. During intervention, the high overlap in PEM data came during non-tutoring days (40%), with tutoring days and over weekly mean scores both exhibiting only 20% of data points overlapping the median baseline score. This PEM data is interpreted to show moderate effectiveness of the tutor training and implementation on Tutor 5's behavioral scores during tutoring days and overall weekly average scores, but questionable effectiveness on non-tutoring days, alone. Maintenance PEM data (50%) showed the intervention's small effect on Tutor 5's CICO weekly averages after the last tutoring session, compared to his baseline median score.

Overall

Overall PEM data showed moderate to large effects of the tutoring intervention on tutors' weekly CICO behavioral point averages on tutoring days, across tutors. Across tutors, PEM during intervention demonstrated large effect, at 92% (SD = 8.4). On non-tutoring days the effects were more variable, with three tutors' CICO scores showing moderate improvement and two tutors' CICO scores showing small, or questionable effects. On non-tutoring days PEM data was found to be 77.4% (SD = 13.6) across tutors, showing moderate effectiveness. Overall, weekly average CICO scores for tutors during intervention phase demonstrated moderate (3 tutors) to large (2 tutors) intervention effectiveness. Across tutors, weekly mean CICO scores showed improvement at the upper end of the moderate effect range (90%; SD = 10.0). During maintenance phase, PEM data showed the tutoring training and intervention to have small to large effects on tutors' weekly CICO mean scores. Across tutors, PEM data showed a mean improvement of 68.8% (SD = 28.6) in CICO scores from baseline to maintenance phase.

Effect Sizes: Pre-/Post-intervention Measures (SRSS)

To further evaluate the effects of the intervention on the tutors' behaviors, a distal measure, the Student Risk Screening Scale (SRSS; Drummond, 1994) was also administered. Pre-intervention SRSS scores were obtained during participant screening. Like tutees' scores, assessment of potential generalized outcomes was evaluated by calculating Hedges' *g* effect sizes for externalizing and internalizing SRSS scores. Table 4.3 shows tutors' pre-/post-intervention scores and EBD risk status.

					EBD risk status			
	SRSS *Pre*		SRSS *Post*		*Pre*		*Post*	
Tutors	Extern.	Intern.	Extern.	Intern.	Extern.	Intern.	Extern.	Intern.
Tutor 1	10	4	8	6	H	M	*M	M
Tutor 2	11	8	7	8	M	M	M	M
Tutor 3	17	13	15	12	H	H	H	H
Tutor 4	13	13	6	12	H	H	*M	H
Tutor 5	15	11	13	9	H	H	H	H
Mean (SD)	13.2 (2.9)	9.8 (3.8)	9.8 (4.0)	9.4 (2.6)				
Effect size (*g*)	Externalizing: *g* = .88 Internalizing: *g* = .11							

Note. SRSS = Student Risk Screening Scale (SRSS; Drummond, 1994); Extern. = externalizing disorder; Intern. = internalizing disorder; SD = standard deviation; H = high risk; M = moderate risk; * = decrease in risk status categorization

Table 4.7: Tutors' Pre-/Post-intervention Scores, Risk Status, and Effect Sizes.

Four of five tutors' pre-intervention scores fell under the classification of high-risk status for externalizing behavioral disorders, and the remaining tutor was found to have a moderate risk status (*M* = 13.2; SD = 2.9). SRSS guidelines suggest that students falling within moderate to high categorization for risk be provided with intervention for challenging behaviors. For internalizing disorders, pre-intervention scores for three tutors fell within the high-risk category, while the remaining two students fell within the moderate-risk classification (*M* = 9.8; SD = 3.8). After the intervention was removed, the SRSS was administered again, and all five tutors were found to have decreased their SRSS risk scores (both internalizing and externalizing) from pre- to post-intervention. Across tutors, externalizing risk scores showed the greatest improvement, with a decrease by more than three points (*M* = 9.8; SD = 4.0), while internalizing risk scores decreased at lower rates or remained at approximately the same levels (*M* = 9.4; SD = 2.6). Furthermore, a change in risk-status was exhibited in two tutors' scores, as they decreased their risk status for externalizing disorders from high to moderate levels. No tutor was observed to have an increase in EBD risk status due to participation in the intervention. Overall, the cross-age tutoring intervention demonstrated large, statistically significant effects in decreasing externalizing disorder risk status (g = 0.88) but questionable effects in improving risk status for internalizing disorders (g = 0.11).

Summary

Research question three, which assessed the extent to which the cross-age tutoring training and implementation of the tutoring program was effective in promoting tutors' (i.e., students with/at-risk for EBD) behavioral performance as measured by CICO point sheets, was evaluated through three sets of analysis: visual analysis, proximal effect sizes (i.e., of visual data), and distal effect sizes (i.e., pre-/post-assessment).

First, in assessing the visual data, a causal relation could not be fully demonstrated between the cross-age tutoring program training and implementation, and the tutors' behavioral performance as measured by CICO point sheets due to the high levels of variability in two tutors' data sets. Although improvements in CICO scores from baseline to intervention phase were seen for a majority of tutors, it was within the maintenance phase that greatest improvements were demonstrated. Behavioral scores on CICO were maintained or increased at, or above baseline and intervention levels. Second, visual data effect sizes (PEM) were calculated and analyzed. Weekly average CICO scores for tutors during intervention phase demonstrated moderate (3 tutors) to large (2 tutors) effects, with more variability during the maintenance phase (PEM = 68.8%; SD = 28.6). Finally, tutors were evaluated for potential changes in risk status for EBD from pre- to post-intervention. Two of five tutors decreased their EBD risk status from high to moderate levels within the externalizing disorders domain, and overall, the cross-age tutoring intervention demonstrated large, statistically significant effects in decreasing externalizing disorder risk status (g = 0.88) but questionable effects in improving risk status for internalizing disorders (g = 0.11).

In sum, through the multiple analyses of tutors' behavioral outcomes, the cross-age tutoring program demonstrated varying effectiveness in promoting CICO behavioral score improvements and changes in risk status for EBD. The training and implementation of the cross-age tutoring program demonstrated small to large, statistically significant effects across tutors to varying degrees, and was most effective in consistently decreasing external disorder risk status.

SOCIAL VALIDITY

The researcher-developed social validity survey contained rating-scale and open-ended questions. The tutor survey contained nine rating-scale questions, the teacher survey contained eight rating-scale questions, and the tutee survey contained seven rating-scale questions; and all surveys contained two open-ended questions. Rating options were 1 = strongly disagree, 2 = disagree, 3= neutral/no change, 4 = agree, 5 = strongly agree. For tutees, the investigator verbally asked the tutees the survey questions and recorded their dictated responses on a paper form. Table 4.8 summarizes the results of the social validity survey for each participant population.

Rating scale questions*	Teacher (Implementer) rating	Tutor mean rating	Tutee mean rating
1: I understood all of the components/requirements/ expectations of the math buddy program.	5.0	4.6	4.5
2: I believe I have the skills needed to be an implementer/tutor/tutee ('math buddy').	5.0	4.4	4.5
3: I believe my math skills/behavior have/has gotten better due to the tutoring program. T: Tutors' positive behaviors and/or social skills have increased due to the tutoring program.	4.0	4.0	4.75
4: I believe that the tutor's/my challenging/problem behaviors have decreased since the start of the tutoring program. T: I believe that challenging/problem behaviors of tutors have decreased due to the tutoring program.	4.0	3.8	N/A
5: The tutoring program was easy for me to do. T: My perception of the tutoring program was that it was relatively easy to implement (i.e., amount of time required, effort, practicality).	3.0	4.6	4.25
6: The tutoring program was worth my time and effort.	5.0	4.2	4.5
7: I believe my tutee/'math buddy' got better at math due to the tutoring program. T: The tutoring program is beneficial for my students.	5.0	4.8	N/A
8: I looked forward to meeting with my 'math buddy' for the tutoring sessions.	N/A	4.75	4.2
9: I would be willing to participate/implement the tutoring program again.	5.0	5.0	5.0

Open-ended questions	Strengths/Positives/Favorite parts	Weaknesses/Challenges/ Suggestions for improvement
Teacher	• Previewing materials and content allowed the younger students to feel more confident in their [general education] classroom • Making time commitments for the sessions with work for all of the students' individual class schedules.	• Scheduling and communication with other teacher [i.e., teacher of tutees]. • Suggestion: Have multiple game designs to keep engagement

Tutors	• Just teaching [him]. He is very smart. • To help the little kids • I like when they learn • Talking and having conversations • Getting to be with [my math buddy] • Calming him down and teaching	• The times of the tutoring [sessions]. • Nothing • Getting off track • Trying to calm [him] down.
Tutees	• Working with my math buddy • Moving the characters • Playing the game and doing the tests • Winning! Picking the game piece	• Spinning the spinner • Nothing to me • Nothing • Losing

Note. * = questions summarized/rephrased for table format; T = teacher question; Rating options were 1 = strongly disagree, 2 = disagree, 3= neutral/no change, 4 = agree, 5 = strongly agree

Table 4.8: Social Validity Survey Results.

In regards to assessing prerequisite skills for participating and implementing the tutoring program, all tutors, tutees, and the implementer all felt that they understood the components/requirements and that they were able to fully participate in the intervention. All participants reported high favorability towards the benefit and effectiveness of the intervention. The teacher reported perceived effectiveness of the tutoring program on increasing tutors' positive behaviors and decreasing negative/challenging behaviors. Tutors scored themselves similarly in these areas but their self-rating of decreasing challenging behaviors due to the program was slightly lower than their teacher's rating, and the closest to the neutral rating (3) of all social validity questions. The highest ratings, across tutors, tutees, and the teacher, came in the area of perception of the intervention's impact on tutees' mathematics skills. Tutors, tutees, and the implementer all perceived the tutoring sessions to be effective in promoting the tutees' knowledge and abilities in mathematics. Additionally, all tutors and tutees reported high agreement with questions regarding the ease of implementation of the program and desire to participate in future incarnations of the program. The teacher reported a neutral rating for ease of implementation, but high agreement in the program's worthwhileness, benefit, and desire to participate in future programming.

The practitioner implementer reported that the strengths of the program were the materials, as they allowed the tutees to preview and learn content that would increase confidence in their classroom. He also stated that a positive outcome of the program was the ability to schedule common times for the tutees and tutors to meet for their sessions. Tutors reported that they enjoyed interacting with their tutee and expressed interest in their learning. Tutees stated that their favorite part of the tutoring program was working with their "math buddy" (i.e., tutor), playing the number line board games, picking and moving the game pieces/tokens, winning the game, and completing the weekly tests.

Tutors reported that the tutoring sessions occasionally interfered with their ability to attend preferred activities. They also stated that the biggest challenge was keeping their tutee on-task and "calm". Tutees frequently stated that there was nothing they disliked about the program or would change, but that spinning the number spinner was difficult at times. The teacher implementer reported that the most difficult part of implementing the program was the scheduling of the tutoring sessions due to the differences in the structure of academic, daily schedules of the tutees and the tutors. He also suggested that future tutoring programs consider developing alternate versions of the board game to potentially assist in retaining engagement and interest of the tutees.

Overall, the program was viewed as effective and beneficial for both sets of student participants (i.e., tutees and tutors) and that the intervention was worth the time and effort to implement. The main challenge was stated to be in scheduling mutually agreeable times for the tutoring sessions based on the unique class schedules of each of the participant groups.

Summary of Chapter

This chapter addressed the results related to the three research questions, and additionally, social validity was assessed. Visual analysis and two sets of effect sizes demonstrated positive effects of the cross-age tutoring intervention on the mathematics performance of kindergarteners (i.e., tutees) at-risk for disability. Effect sizes of visual data (i.e., PEM) showed moderate to large effects, with the mean effect size demonstrating large improvements across tutees' performance on weekly probes, and distal effect sizes related to pre-/post-intervention administration of the TEMA-3 demonstrated statistically significant effects (g = 2.78). Overall, through multiple analyses of the results, the cross-age tutoring program showed effectiveness in promoting at-risk kindergarten students' mathematics performance on both proximal and distal measures of early numeracy knowledge and skills.

Research question two evaluated the tutors' ability to implement instructional procedures with fidelity during intervention, in combination with the ability to increase early numeracy knowledge and skills of tutees. In assessing these two components, results showed that tutors were able to both implement tutoring instructional procedures that they were trained upon, with high rates of fidelity throughout the intervention (across tutors: M = 90.5%; range = 83.6%–95.5%), as well as facilitate large and statistically significant gains in mathematical performance of tutees as assessed by performance on early numeracy measures. The combination of these two findings suggests that students with/at-risk for EBD can effectively serve as cross-age tutors.

The final research question assessed the effectiveness of the tutor training and implementation of the tutoring session on tutors' behavioral scores, as measured by CICO point sheets. The cross-age tutoring program demonstrated effectiveness in promoting CICO behavioral score improvements for a majority of the tutors, and decreases in risk-status for EBD were observed for multiple students. Overall, the training and implementation of the cross-age tutoring program demonstrated varying effects, with small to large, statistically significant effects across tutors, and was most effective in consistently decreasing external disorder risk status.

Finally, all participants of the intervention, including the teacher implementer, viewed the program as effective and beneficial for both tutors and tutees, and was worthwhile to implement. The main challenge was stated to be in scheduling mutually agreeable times for the tutoring sessions to occur, due to the unique differences and needs of the individual class schedules for each participant group (i.e., kindergarteners and upper-elementary students). All participants reported a strong desire to participate in future incarnations of the program.

Chapter 5: Discussion

The purpose of this study was to examine the effects of a cross-age tutoring intervention with number line board games on the mathematical performance of kindergarten students at-risk for mathematics difficulties. Additionally, the tutors implementing the program were identified with, or at-risk for EBD, and were assessed for potential changes in behavioral performance due to the intervention provided tutoring training and implementation of tutoring sessions. Participants in the study included five fifth- and sixth-graders (i.e., the cross-age tutors) and five kindergarten students (i.e., tutees), who formed the five tutor dyads, along with a special education teacher who served as the tutoring program implementer and supervisor of the tutoring sessions.

Young children with or at-risk for mathematics learning disabilities (LD) are often challenged with basic number sense knowledge and skills and these challenges with at an early age can lead to substantial gaps later in life and are associated with detrimental long-term outcomes (Geary et al., 2009; Locuniak & Jordan, 2008; Mazzocco, Feigenson, & Halberda, 2011). Additionally, correlations have been found between early numeracy skills and elementary mathematics achievement (Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Morgan & Farkas, 2009). Furthermore, in the long term, students identified as having deficits in early numeracy knowledge and skills have shown continuing low performance on future measures of mathematics achievement and may struggle to develop the conceptual foundations that will support the learning of more advanced mathematics (Baker, Gersten, & Lee, 2002; Jordan, Kaplan, Locuniak & Ramineni, 2007; Van Luit & Schopman, 2000).

Number knowledge is required to develop mental number lines, which are cognitive constructions of increasing number magnitude from left to right, horizontally. One early numeracy intervention, with a foundation in whole number concepts, is playing number line board games. Empirical research and theoretical frameworks connect playing number line board games to the development of an accurate mental number line (Ansari, 2008; Hubbard, Piazza, Pinel, & Dehaene, 2005). Mental number line formation supports early numeracy knowledge and skills in areas such as number magnitude, number line estimation, and number comparison. Participating in number line board games for short durations has shown effectiveness in promoting early numeracy knowledge and skills in the areas of number magnitude, number line estimation, and number comparison (Laski & Siegler, 2014; Ramani & Siegler, 2008; 2011; Siegler & Ramani, 2009; Whyte & Bull, 2008). Children who engaged in number line board games demonstrated significant improvement in the areas of counting, number comparison, number line estimation, and numerical magnitudes compared to peers who participated in basic early numeracy activities such as identifying numerals and verbal counting (Siegler & Ramani, 2009; Whyte & Bull, 2008).

Two main reasons supported the rationale for utilizing number line board games as the primary instructional materials for this study. First, due to their effectiveness in promoting early numeracy knowledge and skills, and second, because of their perceived practical and feasible requirements in the areas of training time, costs (i.e., materials), and prerequisite mathematics skills (e.g., Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Whyte & Bull, 2008). These features of accessibility may allow non-teachers (i.e., cross-age tutors) to deliver the intervention effectively. Furthermore, in evaluating previous research in the area of cross-age tutoring and students with disabilities, the model has shown promise in not only improving mathematics performance of student receiving tutoring instruction (i.e., tutee), but also facilitating improvements for the tutor(s) in behavioral domains (Watts, Bryant & Carroll, 2018). Henceforth, this study assessed a cross-age tutoring, number line board game's effects on tutees' mathematics performance and tutors' classroom behavior and risk-status for EBD.

A co-occurring multiple baseline design across subjects design was implemented to assess the effects of cross-age tutoring on tutees' mathematics performance (i.e., on proximal and distal measures) and tutors' behavioral ratings (i.e., on daily CICO point sheets) and EBD risk-status. Furthermore, tutors' implementation of instructional components (i.e., fidelity) was also evaluated. Prior to the start of the intervention, the special education teacher/case manager of the students serving as tutors was trained on the intervention procedures, including how to train the tutors prior to the intervention. The tutor training sessions included modeling, guided practice, and role-playing of number line board game procedures, corrective feedback strategies, and positive reinforcement techniques. The tutor training sessions took place one-on-one after a stable baseline was identified for the tutees' weekly TEMI-AC total scores. The intervention phase consisted of three, 25-30-min tutoring sessions per week, over 10-weeks. During these tutoring sessions, the cross-age tutor facilitated the modeling, pacing, and procedures of the number line board games with his tutee. At the end of each week, a weekly progress monitoring probe (i.e., TEMI-AC) was administered to assess mathematics performance.

The intervention's proximal and distal effects were measured for both tutees and tutors. For tutees, the TEMI-AC probes measured proximal, early numeracy knowledge and skill development, and the TEMA-3 measured pre-/post-intervention changes in mathematical performance. For tutors, weekly mean CICO behavioral ratings provided proximal data sets, and the SRSS evaluated any potential changes in risk status for EBD from pre- to post-intervention. Maintenance of effects was assessed at 2- and 4-weeks after the last tutoring session. Additionally, social validity surveys were administered to tutors, tutees, and implementing teacher to evaluate perceptions of the tutoring program.

In this chapter, findings in relation to study's three areas of research are discussed: (1) the intervention's effects on tutees' early numeracy knowledge and skills as measured through proximal and distal measures of mathematics performance; (2) the extent to which students with EBD are able and effective cross-age tutors (i.e., implementing instruction with fidelity and increasing tutees' number sense skills); (3) the intervention's effects on tutors' behavioral performance in academic settings and risk status for EBD as measured through proximal and distal measures, respectively. Additionally, the limitations of the study, future areas of research, and implications for practice are presented.

RESEARCH QUESTION 1

Three sets of analysis were implemented to assess research question one: visual analysis, proximal effect sizes (i.e., of visual data), and distal effect sizes (i.e., pre-/post-assessment). These analyses evaluated the extent to which cross-age tutoring utilizing number line board games improved mathematics performance for kindergarteners at-risk for mathematics disability. A causal relation was demonstrated between the cross-age tutoring program and at-risk kindergarteners' mathematics performance on early numeracy measures. Furthermore, across tutees, maintenance of scores was demonstrated at 2- and 4-weeks after the last tutoring session was attended. Evaluating visual data effect sizes (i.e., PEM), moderate to large effects were identified across tutees' total scores on weekly TEMI-AC probes. Additionally, all tutees demonstrated continual improvement during the maintenance phase, where large effects were present (i.e., compared to baseline levels). Finally, distal measures demonstrated statistically significant effects on tutees' mathematics performance on the TEMA-3 from pre- to post-intervention. Overall, findings suggest cross-age tutoring utilizing number line board games, delivered by tutors with EBD, to be effective in promoting at-risk kindergarten students' mathematics performance on both proximal and distal measures of early numeracy knowledge and skills.

Previous reviews of cross-age instructional models with tutors with EBD have shown consistent, positive effects on mathematics performance of both tutees and tutors (Ryan, Reid, & Epstein, 2004; Watts, Bryant, & Carroll, 2018). Two studies focusing on tutoring early numeracy knowledge and skills showed moderate to large effects on tutee outcomes (d = .68; PEM = 91.7; Watts & Bryant, 2017), and large effects on tutor outcomes (d = 1.0; Holecek, 2012). Similarly, the at-risk kindergarten students in this study showed moderate to large improvements in early numeracy knowledge and skill areas such as counting, number comparison, and number magnitude. In regards to effectiveness, this study's findings are consistent with the literature base showing number line board games to be an effective intervention for young children with mathematics difficulties (Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Whyte & Bull, 2008). Additionally, the findings related to tutees' improved mathematics performance add to the evidence-base suggesting number line board games to be an effective instructional tool for promoting early numeracy knowledge and skills for students with at-risk backgrounds (Ramani & Siegler, 2007; Siegler & Ramani, 2009).

Four previous studies have utilized cross-age tutors with EBD as the interventionists in delivering mathematics instruction to students at-risk or with disabilities (Lazerson, 2005; Holecek, 2012; Maher, 1984; Watts & Bryant, 2017). Across those studies, the evidence base for the model has shown to moderate to large effects on mathematics performance for tutees and has also shown the most consistent, positive outcomes for participants when compared to studies utilizing cross-age tutoring for instruction in other content or skill areas (e.g., reading fluency, spelling) (Ryan, Reid, & Epstein, 2004; Watts, Bryant, & Carroll, 2018). One possible explanation may be that the structured procedural steps for some mathematics skills are more conducive to the cross-age tutoring model than, for example, reading fluency or comprehension skills.

Furthermore, considering the backgrounds of the kindergarteners that qualified to be tutees in this study (i.e., in need of intervention due to mathematics difficulties), four of the five students were identified as coming from low-socioeconomic backgrounds. The positive findings in improved mathematics performance align with previous research showing the number line board games as an effective intervention for students coming from at-risk backgrounds (Ramani & Siegler, 2007). These findings are important because children from low-socioeconomic status families have been shown to have deficits in skills related to number competencies (e.g., counting, adding, subtracting, and comparing magnitudes) compared to peers from families who are more economically advantaged (Arnold, Fisher, Doctoroff, & Dobbs, 2002; Jordan, Kaplan, Olah, & Locuniak, 2006). Additionally, these deficits in early mathematical competencies are strong predictors of later struggles on measures of mathematics achievement (Geary, 2011). The need for effective interventions that promote early numeracy knowledge and skills cannot be understated, but equally important are feasibility considerations for practitioner implementation of these interventions in natural academic settings.

The effectiveness of practitioner implemented number line board games has been infrequently assessed in the literature. Although findings show it to be an effective and engaging instructional support for providing structured practice opportunities and related feedback when implemented by practitioners (e.g., paraprofessionals; Ramani, Siegler, & Hitti, 2012). Considering the limitations of current literature base, this study evaluated the number line board game intervention through a model where the practitioner served as the trainer and supervisor of the tutoring program, while older students were utilized as tutors within a peer-mediated instructional model (i.e., cross-age tutoring) to deliver the game procedures and structured feedback. This arrangement allowed the teacher to observe multiple dyads at the same time, address any challenges when they arose, and informally assess progress of multiple students at the same time. Additionally, this model provided tutees with one-to-one modeling, multiple opportunities to practice, immediate feedback/correction from tutors, and direct, positive reinforcement from an older peer. Most importantly, the instructional arrangement facilitated positive improvements in tutee early numeracy knowledge and skills, demonstrating promising effects as a peer-mediated intervention.

Overall, the findings of this study are consistent with previous, though limited, research showing number line board games to be a promising evidence-based practice for promoting early numeracy knowledge and skills as exhibited by improved mathematics performance on proximal and distal measures (Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Whyte & Bull, 2008). Positive findings related to tutee mathematics outcomes align with research on the intervention's effects for students from low-socioeconomic backgrounds, non-researcher implementation/feasibility, and peer-mediated delivery of tutoring procedures (Ramani, Siegler, & Hitti, 2012; Watts, Bryant, & Carroll, 2018).

RESEARCH QUESTION 2

The effectiveness of students with EBD as cross-age tutors was evaluated through two outcomes: (1) the extent to which tutors implemented tutoring instructional procedures with fidelity during intervention, and (2) the extent to which tutees' early numeracy knowledge and skills were promoted, as measured by mathematics performance during and after the intervention (i.e., proximal and distal outcomes). Findings showed tutors were able to implement tutoring instructional procedures with high rates of fidelity and improve tutees' mathematics performance at statistically significant levels. These findings demonstrate that students with/at-risk for EBD can effectively serve as cross-age tutors.

When students with EBD participate in peer-mediated intervention models they frequently assume or are assigned the role of the tutee and/or share instructional responsibility with a peer (e.g., reverse-role/reciprocal), and rarely are provided the opportunity and responsibility of providing instruction or academic support to other students (Falk & Wehby, 2001; Franca, Kerr, Reitz, & Lambert, 1990; Penno, Frank, & Wacker, 2000). The findings from this study add to a limited literature base in assessing the extent to which students with EBD can effectively serve as cross-age tutors and implement instructional procedures with fidelity throughout the intervention.

Across the small number of studies that measured and reported fidelity of implementation of cross-age tutors with EBD, the rates of implementation were high, ranging from 88% to 97% (Blake et al., 2000; Hamelberg, 1987; Maher, 1984; Watts & Bryant, 2017). Additionally, when fidelity of implementation was measured for cross-age tutors with EBD, the related outcomes for tutees showed moderate to large improvements in academic and non-academic skills, with effects being maintained in each of the studies. This dissertation study found similar results regarding tutors' fidelity of implementation, which was greater than 90%, across tutors. Tutees' mathematics scores on proximal and distal measures exhibited moderate to significant improvements. The findings related to research question two add to the evidence-base suggesting students with EBD are capable and effective cross-age tutors when provided with the appropriate training and supervision.

In regards to training, some improvements may be necessary for future incarnations of instructional models requiring the training of students with EBD as cross-age tutors. Across the 10-weeks of the intervention phase, fidelity of implementation means were high across tutors (i.e., > 90%). These results demonstrate tutors' ability to effectively implement the tutoring procedures, but areas for improvement can be identified when fidelity rates are disaggregated by weekly means (see Table 4.4). Re-training sessions were required (i.e., when weekly fidelity means fell below 80%) and provided for four of five tutors within the first 3-weeks of the intervention phase. There are multiple explanations for these initial, variable levels of implementation.

First, the initial tutor training session may have been insufficient in some regard. Training considerations must be in regard to the effect of training dosage (e.g., intensity/effectiveness of instruction, duration of training sessions), components, and/or personnel. Findings related to increased fidelity rates after re-training sessions were provided suggests that basic modifications to the training protocol could facilitate higher, initial rates of implementation. One recommendation would be to provide more practice opportunities with instructional procedures within training sessions. This may require tutor-training sessions to increase in frequency and duration. This alteration may support the development and maintenance of required tutoring procedures and strategies.

Another barrier to acceptable rates of implementation may be found in the transition of tutoring skills from the training environment into the natural tutoring setting/sessions, as well as the unique differences between the dynamics within those two environments. Evaluating the trend in fidelity levels across tutors as the tutoring program progressed, tutors' rates of implementation consistently increased in level and became stable at or above 90% during later weeks. This finding suggests tutors, and possibly tutees as well, may require time to practice and become familiar with the routine and procedures within the tutoring sessions. Henceforth, tutors may benefit from additional practice/role-playing with other students (e.g., students being trained as tutors) during tutor training. This feature could be provided in addition to initial practice/role-play opportunities with the teacher-trainer, where corrective feedback and reinforcement is provided to the tutor. It is recommended that future incarnations of cross-age models utilizing tutors with EBD be designed with these recommendations in mind.

Overall, evaluating the fidelity of implementation of tutors across the duration of the intervention, findings support the practice of training students with EBD to be effective cross-age tutors. Results align with previous research showing the importance of effective training and supervision of cross-age tutors with EBD (Heron, Welsch, & Goddard, 2003). In sum, students with EBD have demonstrated effectiveness in delivering instructional procedures that support tutee improvements in mathematics performance. The evidence-base for cross-age tutoring as delivered by students with EBD shows promise as an effective instructional arrangement not only for supporting tutees in need of supplemental instruction/practice but also, potentially, for promoting behavioral outcomes for tutors as well.

RESEARCH QUESTION 3

The training and implementation of the cross-age tutoring program demonstrated small to statistically significant effects on tutors' behavioral outcomes to varying and was found to be most effective in consistently decreasing risk for developing external disorders. Evaluation of proximal data found a majority of tutors showed improved performance in CICO scores from baseline to intervention phase, with continual, significant improvements demonstrated during maintenance phase. Effect sizes showed moderate effects for three tutors and large effects for two tutors. Findings related to distal measures showed two of five tutors decreased risk status for EBD from pre-intervention levels, and across tutors, statistically significant decreases in externalizing disorder characteristics were observed (g = 0.88). Given that social-behavioral skills are frequently characterized as deficit areas for individuals with EBD (Landrum, Tankersely, & Kauffman, 2003; Trout, Nordness, Pierce, & Epstein, 2003), utilizing cross-age tutoring shows promise as a possible intervention for addressing these needs.

The cross-age tutoring model has shown greater effects on social-emotional and behavioral skills compared to targeted academic skills, and furthermore, non-academic skills are more readily maintained after the intervention (Bowman-Perrott, Burke, Zhang, & Zaini 2014; Watts, Bryant, & Carroll, 2018). Literature has also shown the model to be more effective for students with disabilities in promoting social skills and reducing disruptive behaviors comapred to increasing academic engagement (Bowman-Perrott et al. 2014).

Within this study, all tutors showed an increase in level of CICO behavioral point averages from baseline to intervention phase. During the intervention phase, tutors' CICO behavioral scores were greater on days when they attended tutoring sessions than on days when tutoring sessions were not scheduled, suggesting a possible relation between implementation of the tutoring procedures and/or interaction with tutees, and their behavioral performance within academic settings. Across tutoring and non-tutoring days, the intervention exhibited moderate effects on tutors' weekly CICO mean scores. These findings suggest the intervention may provide access to collateral outcomes in general academic settings (i.e., social-behavioral improvements). Further research is needed in evaluating the specific features of the intervention (i.e., tutor training or delivery of tutoring sessions) that may be functioning as the mechanism(s) of change.

Evaluating the tutors' behavioral outcomes in regards to lasting effects is more complex. The maintenance of students' CICO behavioral ratings after the intervention was removed can be interpreted two ways. First, students that retained or increased weekly CICO point means during maintenance may have acquired or become familiar with social and/or behavioral skills within the training and implementation of the tutoring intervention that facilitated desired behaviors within general classroom settings at higher rates. Second, for tutors showing decreased CICO means during maintenance, the intervention may have produced consistently higher behavioral ratings during implementation but decreased when tutors were no longer practicing tutoring skills and/or receiving feedback (e.g., from tutees, from supervising teacher) during tutoring sessions/days, which may have impacted the use of similar skills in general academic settings.

Tutor 3's proximal data provides a unique opportunity to evaluate under these terms, as he returned to baseline (i.e., intervention/tutoring was removed) after week three due to his tutee moving out of the district. Additionally, this phase, after the intervention was removed, can also be interpreted as an extended maintenance phase. Considering Tutor 3 did not receive the full dosage of the intervention, relative to the other tutors, his situation allows for the analysis of seven additional weeks of (return to) baseline/maintenance data. Although implemented for a short duration, Tutor 3's results can assist in determining the effects of the intervention. During each of the 3-weeks intervention phase, the tutor's weekly CICO mean scores were higher than any of the weeks of baseline or extended maintenance. These findings demonstrate a possible relation between the intervention and Tutor 3's behavioral improvement during weeks when he was attending tutoring sessions. Two tutors' data during maintenance phase showed significant and lasting effects of the intervention, while three exhibited questionable effects. Although three tutors' maintenance phases CICO scores were lower than the intervention levels, none of the maintenance phase levels were below baseline phase levels. Evaluating maintenance of effects across all five tutors shows the intervention's variable, lasting effects on weekly CICO mean scores.

There are at least two possible explanations for these results. First, for tutors whose data exhibited continual increases, the intervention may have provided the training (e.g., modeling, practicing social/behavioral strategies) and/or feedback/reinforcement (e.g., from tutees, from teachers) that allowed the tutors to learn to use functional social or behavioral skills outside of the tutoring setting (i.e., within their academic classes). Second, tutors whose CICO levels dropped during maintenance phase may have benefited from a component within the tutoring sessions that was either motivating or reinforcing to their behavior during the intervention phase, but when removed (i.e., no longer attending tutoring sessions), may have resulted in classroom behaviors returning to pre-intervention levels.

Robinson, Schofield, and Steers-Wentzell (2005) propose when students with or at-risk for disabilities undertake the role of a cross-age tutor, the large discrepancy between that role and their typical student identity requires a transformation that produces "spillover" effects seen in the form of increases in academic skills, time on-task, classroom behavior, and positive attitudes towards school (Allen & Feldman, 1976; Bierman & Furman, 1981). This explanation connects to *role theory*, which proposes that the behavior of the individual is influenced by the role they inhabit or play (Thomas & Biddle, 1966). In the context of this study, the theory proposes that when students undertake a given role (e.g., cross-age tutor, teacher), they adopt the attitudes or behaviors associated with the assumed role/identity (Turner, 2002). Thus, this theory provides a possible explanation for behavioral improvement during the intervention phase and decreases in those same behaviors after the responsibility/role of being a tutor was removed (i.e., during maintenance phase).

Furthermore, although the proximal data for these tutors was variable during the maintenance phase, across tutors, distal effects were demonstrated in improvements in risk status for EBD. Findings related generalized behavioral improvements and decrease risk-status for EBD align with previous research showing potential benefits of cross-age tutoring on students' non-academic skill areas such as general classroom behavior, on-task behavior, social skills, and peer relationships (Blake et al., 2000; Greenwood, Carta, & Hall, 1988; Greenwood, Delquadri, & Hall, 1989; Gumpel & Frank, 1999; Hogan & Prater, 1993; Maher 1982; 1984). Considering this limited literature base, this intervention shows promise as a peer-mediated instructional model but requires further research in specific areas, one being the infrequent measurement of distal outcomes for tutors. Across studies utilizing cross-age tutoring, there are identified needs for research methodologies designed to directly assess the impact of tutor training and implementation on tutors' academic, social, and behavioral skills in generalized settings (i.e., outside of the tutoring environment) (Watts, Bryant, & Carroll, 2018). Generally, this model shows utility in providing opportunities to practice social and behavioral skills in natural, one-on-one settings requiring engaging and instructing younger students with academic and/or behavioral needs. Considering that these skills are frequently characterized as deficit areas for students with EBD, assessing the effects of these components through tutor outcomes is required in further evaluations of this model.

In sum, addressing the intensive needs of students with EBD requires effective interventions that promote social-emotional-behavioral skills and provide positive learning environments and experiences (Kern, 2015). The cross-age model provides opportunities for tutors with EBD to practice positive social-behavioral skills in a structured academic setting, and additionally, receive social feedback from both peers and teachers. Overall, the findings of this study are consistent with the previous, though limited, literature base showing cross-age tutoring as an effective model for delivering supplemental instruction to students at-risk for future disabilities, as well as possibly providing benefits to tutors with/at-risk for EBD (Blake, Wang, Cartledge, & Gardner, 2000; Gumpel & Frank, 1999; Watts, Bryant, & Carroll, 2018). Further research is required to assess if, and to what extent, the cross-age tutoring model supports the development of deficit skills for functional improvements across settings.

LIMITATIONS

Caution must be used in generalizing the results of brief interventions due to their low external validity (Slavin, 2008). Three limitations need to be considered when interpreting the results of this study. First, the uniqueness of the design implemented to evaluate outcomes for two distinct populations of students with disabilities poses limitations. One challenge in utilizing a co-occurring multiple baseline design across two sets of participant groups is the intertwined intervention schedule, that is, when the independent variable (i.e., tutoring program) is introduced to the tutee, it must also be introduced to the tutor (i.e., tutoring sessions are attended/implemented for the dyad).

Therefore, a rationale and decision must be made as to which of the participant set's baseline data (i.e., tutees or tutors) will be utilized in determining when the intervention will be introduced. For this study, a rationale was provided for utilizing the tutees' TEMI-AC total scores as the primary data set for evaluating the stability of the baseline phase, and therefore, determining when each dyad would be provided the intervention (i.e., provided tutor training and begin attending tutoring sessions). This decision was based on the tutees being the primary recipients of the intervention (i.e., provided supplemental mathematics instruction/practice opportunities within the tutoring sessions), while the tutors were being evaluated for collateral benefits (i.e., generalized effects on behavior). One potential issue with this arrangement occurs when tutee's baseline data quickly stabilizes in trend and level, not providing adequate time for the tutors' baseline data to stabilize before the intervention is introduced.

Second, due to the components, structure, and delivery of the intervention (e.g., training tutors, implementation of number line board games, peer-mediated instruction), it is difficult to assess which component(s) was/were the primary mechanism(s) of change. For example, during tutoring sessions, the tutees were exposed to modeling, multiple practice opportunities, explicit feedback, number line board game materials/procedures, positive reinforcement, and attention from, and interactions with an older peer. Each of these components has evidence of supporting effective instruction and could have potentially assisted the development of early numeracy knowledge and skills. Therefore, although the intervention resulted in positive outcomes for both tutors and tutees, the mechanism of change cannot be fully identified.

Third, and related to previous limitations, is the interpretation of the intervention's direct effects on tutors' behavioral outcomes. Due to the nature of assessing collateral (i.e., generalized) effects, evidence of a direct or casual relation between the tutoring intervention and tutors' behavioral improvements (i.e., CICO scores and EBD risk status) must be interpreted with caution. An example of this limitation can be found within the structure and components of the dependent variable, CICO behavioral point sheets. CICO behavioral categories are general in nature because they are aligned with PBIS school-wide expectations (i.e., be respectful, be responsible, be safe). This supports accessibility and feasibility for practitioners (e.g., scoring multiple students' sheets per period based on the same scoring criteria), but limits the measure's sensitivity to individual student's specific behavioral development. This limitation could be addressed in future research through the implementation of direct observation of operationalized target behaviors. Future considerations should also be made in measuring tutors' behaviors within tutoring settings and natural classroom settings (i.e., academic classes; generalized behaviors). These data would provide a comprehensive assessment of potential behavior changes. A casual relation can be interpreted with increased reliability and validity when data collection measures are implemented that utilize direction observation techniques and focus on target behaviors directly related to deficit skill areas.

In sum, potential limitations must be considered when interpreting the results of this study. Challenges have been identified in the areas of the study's design as it relates to establishing a stable baseline for tutors, identifying direct or casual relations through collateral outcomes, and dependent measures assessing tutors' general behavior change in general academic settings. The interpretation of the findings presented in this study must be evaluated with these limitations in mind.

FUTURE RESEARCH

Research is limited in the area of cross-age tutoring models containing tutors with

EBD (Spencer, 2006; Spencer, Simpson, & Oatis, 2009; Watts, Bryant, & Carroll, 2018). This study was designed to address many of the holes in the evidence-base regarding practitioner implementation, generalized behavioral outcomes for tutors, and fidelity of implementation, among others. And yet, there is still much to be covered in the evaluation of this model in terms of replication/external validity, component analysis/dosage, measurement of tutor outcomes, implementers, methodological rigor, and potential benefits for other at-risk populations.

First and foremost, due to the innovative nature of this study, and considering the limited research, there is a need for replication of these methods and procedures in order to assess the external validity. Additionally, considering the instructional components of this intervention (i.e., modeling, role-playing, feedback, and positive reinforcement) are similar to other peer-mediated instructional models (e.g., Peer Assisted Learning Strategies; Fuchs & Fuchs, 2005; classwide peer tutoring; Greenwood, Delquadri, & Hall, 1989), these findings could provide an opportunity for the creation of a more standardized, systematic cross-age tutor training and implementation procedures. Considerations for evaluating components of the cross-age tutoring model should include identifying the maximum number of tutees (i.e., per group) that can be effectively provided instruction by a single tutor as well as the effect of the dosage (i.e., frequency and duration) on participant outcomes.

Conducting a component analysis may also assist in identifying the mechanisms of change that promote target skills for the tutees and/or tutors. Determining the minimum dosage required to improve target skills would also assist in the standardization of the intervention's procedures as well as identify the model's utility for specific student needs. Additionally, considering the cross-age model's utility in providing supplemental instruction, identifying the effects of the intervention's duration on student outcomes may be beneficial for supporting its use as a booster/supplemental instructional practice for students at-risk. Lane, Pollack, and Sher (1972) conducted a cross-age tutoring intervention that lasted 7 months and yielded consistent, positive improvements across both academic and behavioral skills. These findings related to frequency and duration of supplemental instruction align with previous research showing intensifying intervention dosage better meets the needs of students with disabilities (Bryant et al., 2011; Vaughn et al., 2012). Furthermore, a majority of previous cross-age tutoring studies with students with EBD reported intervention phases of 10 weeks or less, signaling a need for further research in this area (Watts, Bryant, & Carroll, 2018).

Moving forward, further research is also recommended in evaluating the model's effectiveness in other academic and social-behavioral domains. This study has shown cross-age tutoring with tutors with EBD to be effective in delivering number line board game materials and procedures to students at-risk for mathematics disabilities. Bowman-Perrott and colleagues (2013) examined peer tutoring effects on academic skills and found the model to be highly effective for students with EBD–obtaining greater benefits from the model than other disability types. Related to expanding the use of cross-age tutoring is the need for research in determining its effectiveness for other population of students with/at-risk for disabilities (i.e., as tutees and tutors). This study found potential behavioral benefits for students with/at-risk for EBD. Future research should assess potential benefits for student populations with similar, but specific social-emotional-behavioral needs, such as autism, intellectual disabilities, and attention deficit hyperactivity disorder (ADHD; other health impairment). Overall, external validity is required in evaluating the model's effectiveness in producing reliable outcomes in natural learning environments (i.e., practitioner implementation).

Along with evaluating the model's effectiveness in supporting different student populations and content area/behavioral skills, practitioner populations must also be considered in determining possible benefits and utility of the model. Previous research has shown paraprofessionals as effective implementers and supervisors of number line board game booster lessons (Ramani, Siegler, & Hitti, 2012). This study has further shown the feasibility and effective implementation of number line board games by cross-age tutors with EBD. Additionally, previous research has shown cross-age tutoring to be more effective than same-age or reciprocal tutoring for students with EBD (Perrott, Burke, Zhang, & Zaini, 2014). Similar findings were obtained through a meta-analysis of tutoring models for literacy instruction, where cross-age tutoring was found to be more effective than adult tutoring and computer-based tutoring, especially when students with disabilities served as tutors (Jun, Ramirez, & Cumming, 2010). Alternatively, differential effects of cross-age instruction by tutors with EBD versus teacher-led instruction has been infrequently evaluated in previous literature (e.g., Maher, 1982; 1984; Tops & Osguthorpe, 1987) with results that are difficult to generalize due to issues in meeting quality indicators for rigorous research methods. A future line of research could evaluate potential differences in levels of engagement of the tutee in relation to type of instructor (e.g., teacher, cross-age tutor with/without disability, paraprofessional).

Finally, considering the variable quality of previous studies utilizing students with EBD as cross-age tutors (Watts, Bryant, & Carroll, 2018), it is recommended that any of the previously outlined research considerations be designed in alignment with the most current evidence-based research standards (i.e., Council for Exceptional Children, 2014). One specific example of how this study would increase rigor in future replications would be to assess the extent to which the general education teachers and special education paraprofessional who scored the tutors' CICO behavioral point sheets were blind to the tutoring program components. Brief surveys during and after the intervention phase could be used in determining the extent to which these practitioners were aware of the intervention. Additionally, it is recommended that rigorous procedures are also utilized in the reporting of study results to determine accurate relations between interventions and participant outcomes (Conn & Chan, 2015; Conn & Groves, 2011).

In sum, further evaluation is required in assessing the external validity (i.e., replication of effects), mechanisms of change (i.e., component analysis), generalized effects (i.e., other content and/or skill areas), and potential benefits for students with/at-risk for other disability types (e.g., autism, ADHD, intellectual disabilities). Additionally, recommendations are presented for increased rigor and reporting standards for future cross-age tutoring studies utilizing students with EBD as tutors.

IMPLICATIONS FOR PRACTICE

Based on this study's findings, there are multiple implications for practice. First, special education classrooms show a continuing demand for effective instructional techniques and arrangements that meet the needs of students with disabilities. Barriers to certain instructional models and interventions become compounded when considerations must be made for cost and/or feasibility (e.g., staff/personnel requirements, necessary training) (Bettini, Kimerling, Park, & Murphy, 2015; Blanton, Sindelar, & Correa, 2006; Brownell, Ross, Colon, & McCallum, 2005; Greenwood, Carta, & Hall, 1988). Cross-age tutoring has shown promising evidence of effectiveness and feasibility for practitioner implementation. Additionally, the model may be able to address the intensive needs of students with/at-risk for learning disabilities while also providing tutors with EBD opportunities to practice and develop social and behavioral skills in an academic context.

The social validity findings were positive, with the teacher implementer/supervisor of the cross-age tutoring program perceiving the intervention as practical and beneficial for all student participants. The main area of challenge was scheduling common times for two student populations (i.e., tutees and tutors; kindergarteners and upper elementary students) to meet for their individual tutoring sessions. Considering all participating students were located within the same school, the logistical barriers were mainly found in scheduling days and times to meet. These logistical concerns must be considered when implementing any cross-age tutoring program, especially when utilizing student tutors from upper grade levels, who typically follow more structured schedules (e.g., periods), allowing less flexibility.

Returning to the social validity of the intervention, all tutors and tutees perceived the program as effective in promoting mathematics skills for tutees and positive behavioral developments for tutors. Additionally, all students expressed high levels of interest in participating in future programs. Overall, social validity outcomes align with previous research showing practitioners' positive perceptions of the program's effectiveness and benefits (Blake, Wang, Cartledge, & Gardner (2000; Cochran, Feng, Cartledge, & Hamilton, 1993; Gumpel & Frank, 1999; Lazerson, 2005).

Second, related to the tutees' outcomes, this model has shown effectiveness in the delivery of number line board game instruction to support students at-risk for mathematics difficulties. Findings align with previous research showing number line board games to be an effective intervention for at-risk kindergarteners from low socioeconomic backgrounds (Laski & Siegler, 2014; Ramani & Siegler, 2008; Siegler & Ramani, 2008, 2009; Whyte & Bull, 2008). Considering that low socioeconomic status appears to be a key factor related to early struggles in mathematical development (Jordan, Kaplan, Ramineni, & Locuniak, 2009), this program shows promise as an effective intervention for providing modeling, extra practice opportunities, and explicit feedback on early numeracy knowledge and skills. This study has also shown older students with

EBD to be capable tutors when provided with the appropriate training and supervision. This instructional arrangement may allow practitioners to utilize variations of this model to simultaneously address students with more intensive needs. For example, a special educator could provide intensive intervention to a small group of students at the same time cross-age tutoring sessions are taking place with a paraprofessional serving in a supervision/supporting role. Henceforth, alternative instructional arrangements may be available for meeting the needs of multiple populations of students with/at-risk for disabilities due to the flexibility of cross-age tutors who are effective, accessible (i.e., available within the school), and require no additional costs.

The implications of this study's findings are promising, considering the evidence showing teachers and students are more likely to continue using the practice with fidelity when they perceive it to be effective or beneficial (Zimmerman & Risemberg, 1997). Furthermore, this intervention aligns with the development of special educator preparation programming that teaches and supports proactive instead of reactive practices (e.g., timeout, removal/exclusion from general education setting) when working with students with EBD (Oliver & Reschly, 2010).

SUMMARY

The purpose of this dissertation study was to investigate the effects of a cross-age tutoring model with tutors with EBD delivering instruction, through number line board games, to children (i.e., tutees) with mathematics difficulties. The study utilized a multiple baseline design across two participant populations (i.e., tutees and tutors) to determine the effects of the intervention on early numeracy performance of tutees and classroom behavior of tutors. Two sets of concurrent multiple baseline designs were implemented–one set for the tutee participants and one set for the tutor participants. The tutees' dependent variable was their mathematical performance (i.e., total scores) on the Texas Early Mathematics Inventory–Aim Checks (TEMI–AC; University of Texas System/Texas Education Agency, 2009), and the tutors' dependent variable was their weekly averaged scores on their Check-in/Check-out behavioral point sheets. The independent variable for the tutees was attending the cross-age tutoring sessions in which they participated in number line board games for 20-25-min per day, 3 days per week, over 10-weeks. The independent variable for the tutors contained two components: (1) tutor training sessions where they received instruction on tutoring skills, number line board game procedures, and positive behavioral reinforcement strategies; and (2) implementation of the cross-age tutoring program through individual tutoring sessions with their tutee.

Based on the visual analysis and proximal data effect sizes, a causal relation was demonstrated between the cross-age tutoring program containing number line board games and the mathematics performance of at-risk kindergarteners on early numeracy measures. These findings suggest cross-age tutoring with number line board games to be an effective intervention for promoting mathematics performance of kindergarten students at-risk. Furthermore, this intervention model shows effectiveness in supporting early numeracy knowledge and skills that are retained and continue to development after the intervention is removed. Findings are consistent with previous, though limited, research showing number line board games to be a promising evidence-based practice for promoting early numeracy knowledge and skills (Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Whyte & Bull, 2008). More specifically, these findings align with previous evidence showing the intervention's effectiveness in improving mathematics performance of students from low socioeconomic backgrounds, who are frequently at-risk for developing mathematics disabilities (Jordan, Kaplan, Ramineni, & Locuniak, 2009; Ramani, Siegler, & Hitti, 2012).

This study also adds to a limited literature base evaluating the fidelity of implementation of peer-mediated interventions when students with EBD serve as tutors (Blake et al., 2000; Hamelberg, 1987; Maher, 1984; Watts, Bryant, & Carroll, 2018). Findings from this study show tutors with EBD to be capable of both implementing tutoring instructional procedures with high rates of fidelity, and producing improvements in tutees' mathematical performance. The combination of these two findings demonstrates the ability of students with/at-risk for EBD to effectively serve as cross-age tutors. The evidence-base for the cross-age tutoring model as delivered by students with EBD shows promise as an effective instructional arrangement not only for supporting tutees but also, potentially, for promoting behavioral outcomes for tutors.

The cross-age model provides opportunities for students with EBD, as tutors, to practice positive social-behavioral skills in a structured academic setting, and additionally, receive social feedback from both peers and teachers. This study identified possible benefits in the area of improving (i.e., lowering) risk-status for the development external disorders, but proximal behavioral outcomes were more variable across tutors. Further research is required to assess the extent to which cross-age tutoring supports the development of deficit skills for functional improvements across settings for students with EBD (Kern, 2015).

Limitations need to be considered in regards to interpretation and generalization of findings related to tutor outcomes. Challenges have been identified and discussed in relation to the design's evaluation of outcomes for two distinct populations of students with disabilities who are intertwined in the same intervention. Potential limitations include the establishment stable baseline phase data for tutors and conservative interpretation of maintenance phase data. These limitations form a foundation for future research in the area of cross-age models utilizing tutors with EBD.

Research recommendations have outlined for increasing rigor and reporting standards in the development of an expanding literature base in this area. Further research is required in evaluating the model's external validity (i.e., replication of effects), mechanisms of change (i.e., component analysis), generalized effects (i.e., other content and/or skill areas), and potential benefits for students with/at-risk for other disability types (e.g., autism, ADHD, intellectual disabilities). Acknowledging the research needs, this study's findings show the intervention to be a promising, practitioner-implemented model for promoting early numeracy knowledge and skills for tutees at-risk and social-behavioral skills for tutors with EBD. Based on the findings related to effectiveness and social validity, there are multiple implications for practice. Positive perceptions related to tutee and tutor outcomes (i.e., effectiveness of the model), and the practicality and feasibility of practitioner implementation suggest that this model can be utilized effectively in natural educational environments.

In sum, this intervention's findings align with previous research (e.g., Blake et al., 2000; Hamelberg, 1987; Maher, 1984; Ramani, Siegler, & Hitti, 2012; Siegler & Ramani, 2009; Whyte & Bull, 2008) and show the instructional model to be effective for students at-risk for mathematics disabilities. Additionally, the cross-age tutoring model may be beneficial for certain students with/at-risk for EBD, as improvements in general behavioral functioning in classroom settings, and overall risk status for EBD, may be available through the training and implementation of tutoring procedures, skills, and responsibilities. This study lays the groundwork for a future line of research in studying the effectiveness, feasibility/social validity, and related outcomes of an academic cross-age tutoring intervention delivered by students with EBD.

APPENDIX A

TEMI-AC Samples – Weekly Probes (Tutees)

TEMI K

Texas Early Mathematics Inventories

Aim Check

FORM A

Student Name

____

MC ______
NID ______
NS ______
QR ______

Total ____

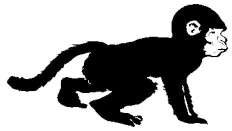
Kindergarten TEMI-Aim Check, Form A

13	4	2	10	1	2	12	4
4	18	3	3	3	1	14	1
10	7	6	13	5	4	2	12
5	15	3	11	7	7	4	8
14	4	15	3	7	5	4	6
9	5	10	9	7	9	14	14
6	8	14	7	11	9	10	12
2	5	13	10	15	13	11	15

19	19	12	11	7	12	9	8
13	11	7	6	12	14	13	12
20	5	16	4	19	4	5	17
20	4	19	6	8	17	18	7
16	6	20	10	10	18	20	9
12	18	13	20	19	12	9	16
20	11	17	11	16	19	18	16
19	20	17	16	20	17	17	18

STOP

1



Number Identification

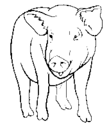
	1	2	3	4
	2	3	4	5
	1	2	3	4
	1	2	3	4
	1	2	3	4
	3	4	5	6
	5	6	7	8

1

	3	4	5	6
	5	6	7	8
	11	12	13	14
	9	10	11	12
	14	15	16	17
	10	11	12	13
	12	13	14	15

1

STOP



Number Sequences

7 8 ___	
6	9
11	10

5 6 ___	
7	17
9	4

4 5 ___	
6	3
14	13

___ 5 6	
2	7
3	4

1 ___ 3	
4	2
0	5

4 ___ 6	
3	7
17	5

8 9 ___	
7	11
10	5

5 ___ 7	
10	6
8	4

___ 1 2	
0	2
4	9

1 2 ___	
4	3
0	5

6 ___ 8	
14	10
9	7

___ 10 11	
1	9
5	12

0 ___ 2

7	3
1	10

___ 1 2

3	4
0	11

0 1 ___

11	2
5	6

6 ___ 8

10	9
5	7

1 2 ___

5	0
3	13

___ 11 12

13	9
10	6

___ 9 10

13	11
8	20

4 ___ 6

7	5
3	10

10 ___ 12

13	11
15	9

13 14 ___

12	5
15	9

___ 15 16

14	18
17	4

13 ___ 15

5	16
14	12

STOP

1

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6	4	5	6	4	5	6

1

STOP

TEMI K

Kindergarten TEMI-Aim Check,
Form A

APPENDIX B

Student Risk Screening Scale – Pre/Post Measure (Tutors)

Student Risk Screening Scale (SRSS) &
Student Internalizing Behavior Screening Scale (SIBSS)

Teacher name:

0 = Never	For the SRSS and SIBBS separately	9–21 indicates high risk	
1 = Occasionally		4–8 indicates moderate risk	
2 = Sometimes		0–3 indicates low risk	
3 = Frequently			

Use the above scale to rate each item for each student.

	Student Risk Screening Scale (SRSS) Externalizing Behaviors								Student Internalizing Behavior Screening Scale (SIBSS)Internalizing Behaviors							
Student Name	Steal	Lie, Cheat, Sneak	Behavior Problem	Peer Rejection	Low Academic Achievement	Negative Attitude	Aggressive Behavior	**Total**	Nervous or Fearful	Bullied by Peers	Spends Time Alone	Low Academic Achievement	Withdrawn	Sad or Unhappy	Complains about Being Sick or Hurt	**Total**

APPENDIX C

Check-in/Check-out Behavioral Point Sheet – Probes (Tutors)

DAILY BEHAVIORAL REPORT CARD

Name: ____________________

Date: ___/___/___

Points Received ______

Points Possible ______

TUTORING TODAY?

YES = ☐

NO = ☐

3 – No redirections/reminders needed
2 – One redirection/reminder given
1 – Two redirections/reminders given
0 – Three or more redirections/reminders given

**Tutoring Period: (Mark with X)*							
Target Behaviors	**[CLASS:]**	**[CLASS:]**	**[CLASS:]**	**[CLASS:]**	**[CLASS:]**	**[CLASS:]**	**[CLASS:]**
Be Respectful -Raise hand -Appropriate language	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0
Be Responsible -On-task -Prepared for class	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0
Be Safe -Hands/feet to self -Use materials as intended	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0	3 2 1 0
Totals =	__ /9	__ /9	__ /9	__ /9	__ /9	__ /9	__ /9

APPENDIX D

Social Validity Survey for Tutors/Tutees

Student Name ___________________________ Date ____________________________________

For each statement, circle one number that best describes how you feel about the tutoring program.

1. I understood all of the elements of the tutoring program.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

2. I believe I have the skills needed to participate as a tutor.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

3. My problem behaviors have decreased in my classes since becoming a tutor.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

4. My classroom behaviors/social skills have improved as a result of the this program.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

5. I believe this tutoring program was relatively easy for me to do.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

6. This tutoring program was worth my time and effort.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

My favorite part of the tutoring program was:

My least favorite part of the tutoring program was:

I would volunteer to participate in this program again: YES or NO

Social Validity Survey for Teachers

Teacher ________________________________ Grade Level: _____________ Date ___________

For each statement, circle one number that best describes how you feel about the tutoring program for your student(s).

1. Academic challenges (in mathematics) have decreased since the implementation of the tutoring program.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

2. Number sense/math skills have increased as a result of the implementation of the intervention.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

3. Social skills/positive behaviors have increased as a result of the implementation of the tutoring program.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

4. My perception of the implementation of the tutoring program was that it was relatively easy (e.g. amount of time/effort/scheduling of time; practical) to implement.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

5. Implementing the tutoring program for this student was worth the time and effort.

Strongly Disagree	Disagree	Neutral/No change	Agree	Strongly Agree
1	2	3	4	5

6. Would you be willing to implement this program in the future: **YES / NO**

Positive aspects of the program:

Challenges of the program:

Any other feedback/observations/recommendations:

Appendix E

TEMI-AC Fidelity Check

Magnitude Comparisons

Tells students purpose for testing.

Reminds students to mark an X if students change their answer and to work until they hear, "Stop."

Sets timer for 2 minutes.

"Turn to shoes… Eyes on me… Hold pencil up."

Provides directions for MC. Tell students to begin and starts timer.

Prompts students during testing.

Stops after 2 minutes; tells students to put pencils down.

________ of 7

Number Sequences

Sets timer for 2 minutes.

"Turn to monkey… Eyes on me… Hold pencil up."

Provides directions for NS. Tell students to begin and starts timer.

Prompts students during testing.

Stops after 2 minutes; tells students to put pencils down.

________ of 5

Place Value

Sets timer for 2 minutes.

"Turn to pig… Eyes on me… Hold pencil up."

Provides directions for PV. Tell students to begin and starts timer.

Prompts students during testing.

Stops after 2 minutes; tells students to put pencils down.

________ of 5

Addition/Subtraction Combinations

Sets timer for 2 minutes.

"Turn to mouse… Eyes on me… Hold pencil up."

Provides directions for ASC. Tell students to begin and starts timer.

Prompts students during testing.

Stops after 2 minutes; tells students to put pencils down.

________ of 5

Total:________/ 22 x 100 =_________%

APPENDIX F

TUTOR TRAINING – Fidelity of Implementation

Training Provided by: __________________ Researcher or Teacher

Fidelity Check Conducted by: _________________ Date: _________

1. The trainer provided rationale for the tutoring system.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

2. The trainer displayed and explained the board game.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

3. The trainer explained the game skills/rules to the students.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

4. The trainer modeled the game skills/rules for the students.
1. - Never implemented 2 - Partially implemented 3 - Fully implemented

5. The trainer role played the game skills/rules for the students with second trainer or trained student.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

6. The trainer had the children role play the skills with a trainer.
1. - Never implemented 2 - Partially implemented 3 - Fully implemented

7. The trainer had the children role play the skills with each other.
1- Never implemented 2 - Partially implemented 3 - Fully implemented

8. The trainer explained the PBS skills to the students.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

9. The trainer modeled the PBS skills for the students.
1. - Never implemented 2 - Partially implemented 3 - Fully implemented

10. The trainer role played the PBS skills for the students with second trainer or trained student.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

11. Reminds the students of any previous skills (board game rules) including the new skill.

1. - Never implemented 2 - Partially implemented 3 - Fully implemented

12. The trainer reviews the key skills of the training session.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

---

DURING ROLE PLAYING:

1. The trainer/teacher provides prompts for students to follow skills taught.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

2. The trainer/teacher provides verbal praise when students are following the skills taught.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

3. The trainer/teacher redirects students if they engage in any problem behavior.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

4. The trainer/teacher re-teaches the students of a skill if a student is having difficulty.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

5. The teacher/trainer provides a 1-2 minute warning until role play is over.
1 - Never implemented 2 - Partially implemented 3 - Fully implemented

Total:________/ 51 x 100 =_________%

APPENDIX G

Tutoring Sessions: Fidelity Checklist

Observer: ____________________ Date: __/__/__

Please complete the fidelity checklist below at rotating between tutors at **30-second intervals**.

***Scoring:** **Y = Yes/Observed** **N = No/Not Observed** **N/A = Not applicable at time of observation**

Instruction:	**Tutor 1 Initials:**									**Tutor 2 Initials:**									**Tutor 3 Initials:**									**Tutor 4 Initials:**								
The tutor . . . **Interval #:**	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
• **Follows game procedures** (takes turns, etc.)																																				
• Provides immediate, **corrective, appropriate feedback** as needed																																				
• **Models** counting strategy, when appropriate																																				
• **Provides practice opportunities** for the math buddy (tutee) to use counting strategies																																				
• **Transitions** efficiently from one game to the next																																				
• **Provides positive reinforcement** through verbal statements or behaviors (high fives)																																				
• **Keeps tutee on-task** while playing game and/or taking math quiz																																				
• **Totals:** • Total Ys / Total Ys + Ns	/									/									/									/								

TOTALS FOR THE SESSION:

Calculate Fidelity %: # of Ys / # Ys + # Ns	**Tutor 1 Initials:**	**Tutor 2 Initials:**	**Tutor 3 Initials:**	**Tutor 4 Initials:**
	_____ %	_____ %	_____ %	_____ %

References

Allen, V.L. (1976). *Children as teachers*. New York: Academic Press. doi:10.1016/c2013-0-07126-5

Allen, V. L., and Feldman, R. S. (1973). Learning through tutoring: Low achieving children as tutors. *Journal of Experimental Education, 42*: 1–5.

Ainley, J. (1990). Playing games and learning mathematics. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education: International perspectives* (pp. 84 –91). Hillsdale, NJ: Erlbaum.

Andres, M., Di Luca, S., & Pesenti, M. (2008). Finger counting: The missing tool? *Behavioral and Brain Sciences, 31*, 642-643.

Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math early numeracy interventions. *International Journal of Early Childhood Special Education, 6*(1), 25-54. 42

Askew, M., Bibby, T., & Brown, M. (2001). *Raising attainment in primary number sense: From counting to strategy*. London, England: BEAM Education.

Aunio, P., Hautamäki, J., Sajaniemi, N., & Van Luit, J. E. H. (2009). Early numeracy in low-performing young children. *British Educational Research Journal, 35*, 25-46

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

Baker, S., Gersten, R., & Lee, D.-S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal, 103*, 51–73.

Ball, D. L. (1992). Magical hopes: Manipulatives and the reform of math education. *American Educator, 16*, 14 –18.

Baroody, A. J., Eiland, M., & Thompson, B. (2009). Fostering at-risk preschoolers' number sense. *Early Education & Development, 20*, 80-128.

Barrett, S., Bradshaw, C. P., & Lewis-Palmer, T. (2008). Maryland state-wide PBIS initiative: Systems, evaluation, and next steps. *Journal of Positive Behavior Interventions, 10*, 105–114.

Bettini, E., Kimerling, J., Park, Y., & Murphy, K. M. (2015). Responsibilities and instructional time: Relationships identified by teachers in self-contained classes for students with emotional and behavioral disabilities. *Preventing School Failure: Alternative Education for Children and Youth*, *59*(3), 121–128. doi:10.1080/1045988X.2013.859561

Blanton, L. P., Sindelar, P. T., & Correa, V. I. (2006). Models and measures of beginning teacher quality. *Journal of Special Education, 40*, 115–127. doi:10.1177/00224669060400020201

Blake, C., Wang, W. Q., Cartledge, G., & Gardner, R. (2000). Middle school students with serious emotional disturbance serve as social skills trainers and reinforcers for peers with SED. *Behavioral Disorders, 25*(4), 280–298. Retrieved from http://www.jstor.org.ezproxy.lib.utexas.edu/stable/43153755

Bierman, K. L., and Furman, W. (1981). Effects of role and assignment rationale on attitudes formed during peer tutoring. *Journal or Educational Psychology, 73*, 33–40.

Bowman-Perrott, L., Burke, M. D., Zhang, N., & Zaini, S. (2014). Direct and collateral effects of peer tutoring on social and behavioral outcomes: A meta-analysis of single-case research. *School Psychology Review*, *43*(3), 260–285.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, *42*(1), 39–55.

Brownell, M. T., Ross, D. D., Colon, E. P., & McCallum, C. L. (2005). Critical features of special education teacher preparation: A comparison with general teacher education. *Journal of Special Education, 38*(4), 242–252. doi:10.1177/00224669050380040601

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannensteil, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children, 78*, 7–23. doi:10.1177/001440291107800101

Calhoon, M. B. (2005). Effects of a peer-mediated phonological skill and reading comprehension program on reading skill acquisition for middle school students with reading disabilities. *Journal of Learning Disabilities, 38*, 424–433. doi:10.1177/00222194050380050501

Carruthers, E., & Worthington, M. (2004). Young children exploring early calculation. *Mathematics Teaching, 187*, 30-34.

Chen, C.-W., & Ma, H.-H. (2007). Effects of treatment on disruptive behaviors: A quantitative synthesis of single-subject researches using the PEM approach. *The Behavior Analyst Today*, *8*(4), 380–397. doi:10.1037/h0100629

Clarke, B., Doabler, C. T., Kosty, D., Nelson, E. K., Smolkowski, K., Fien, H., & Turtura, J. (2017). Testing the efficacy of a kindergarten mathematics intervention by small group size. *AERA Open*, *3*(2), 2332858417706899. doi:10.1177/2332858417706899

Clements, D. H., & Sarama, J. (2011). Early childhood mathematics intervention. Science, 333, 968-970.

Cochran, L., Feng, H., Cartledge, G., & Hamilton, S. (1993). The effects of cross-age tutoring on the academic achievement, social behaviors, and self-perceptions of low-achieving African-American males with behavioral disorders. *Behavioral Disorders*, *18*(4), 292–302.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic

Press.

Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *The Journal of Special Education*, *19*(4), 483–492. doi:10.1177/002246698501900410

Council for Exceptional Children (2014). *Council for Exceptional Children standards for evidence-based practices in special education*. Retrieved from http://www.cec.sped.org/~/media/Files/Standards/Evidence%20based%20Practices%20and%20Practice/EBP%20FINAL.pdf

Damon, W., & Phelps, E. (1989). Strategic uses of peer learning in children's education. In T. J. Berndt & G. W. Ladd (Eds.), *Peer relationships in child development* (135–157). New York: John Wiley and Sons.

Desoete, A., & Gregoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences*, *16*(4), 351–367.

Doabler, C. T., Clarke, B., Stoolmiller, M., Kosty, D., Fien, F., Smolkowski, K., & Baker, S. K. (2017). Treatment intensity and instructional interactions: Exploring the black box of a Tier 2 mathematics intervention. Remedial and Special Education. Advance online publication. doi:10.1177/0741932516654219

Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428–1446. doi:10.1037/0012-1649.43.6.1428

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

Falk, K. B., & Wehby, J. H. (2001). The effects of peer-assisted learning strategies on the beginning reading skills of young children with emotional or behavioral disorders. *Behavioral Disorders, 26*, 344–359.

Fantuzzo, J.W., Davis, G.Y., & Ginsburg, M.D. (1995). Effects of parent involvement in isolation or in combination with peer tutoring on student self-concept and mathematics achievement. *Journal of Educational Psychology, 87*(2), 272–281. doi:10.1037//0022-0663.87.2.272

Fantuzzo, J.W., King, J.A., & Heller, L.R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology, 84*(3), 331–339. doi:10.1037//0022-0663.84.3.331

Franca, V. M., Kerr, M. M., Reitz, A. L., & Lambert, D. (1990). Peer tutoring among behaviorally disordered students: Academic and social benefits to tutor and tutee. *Education and Treatment of Children*, *13*(2), 109–28.

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *The Journal of Special Education*, *39*(1), 34–44. doi:10.1177/00224669050390010401

Frank, A. R., & Sitlington, P. L. (1997). Young adults with behavioral disorders—Before and after IDEA. *Behavioral Disorders*, *23*(1), 40–56. doi:10.1177/019874299702300103

Fuchs, L. S., Fuchs, D., & Hollenbeck, K. N. (2007). Extending responsiveness to intervention to mathematics at first and third grades. *Learning Disabilities Research & Practice, 22*, 13–24. doi:10.1111/j.1540-5826.2007.00227.x

Fuchs, L. S., Fuchs, D., & Kazdan, S. (1999). Effects of peer assisted learning strategies on high school students with serious reading problems. *Remedial and Special Education, 20*, 309–318. doi:10.1177/074193259902000507

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D.

(2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520–1533. doi:10.1111/j.1467-8624.2010.01489.x

Gaustad, J. (1993). Peer and cross-age tutoring. *ERIC Digest 79*. (ERIC Document Reproduction Service No. ED 354 608).

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345–362. doi:10.1037/0033-2909.114.2.345

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*, 4–15.

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K., & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development, 34*, 411–429.

Gelman, R. & Gallistel, C. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*, 1202-1242.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 293–304.

Ginsburg, H. P. and Baroody, A. J. (2003). *Test of Early Mathematics Ability–Third edition (TEMA-3)*. Austin, TX: PRO-ED.

Greenwood, C. R., Carta, J. J., & Hall, R. V. (1988). The use of peer tutoring strategies in classroom management and educational instruction. *School Psychology Review, 17*(2), 258–275.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, *81*(3), 371–383. doi:10.1037/0022-0663.81.3.371

Gresham, F. M., Lane, K. L., Macmillan, D. L., & Bocian, K. M. (1999). Social and academic profiles of externalizing and internalizing groups: Risk factors for emotional and behavioral disorders. *Behavioral Disorders*, *24*(3), 231–245. doi:10.1177/019874299902400303

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

Gumpel, T., & Frank, R. (1999). An expansion of the peer-tutoring model: Cross-age peer-tutoring of social skills among socially rejected boys. *Journal of Applied Behavioral Analysis, 32,* 115–118. doi:10.1901/jaba.1999.32-115

Gunter, P. L., Jack, S. L., Shores, R. E., Carrell, D. E., & Flowers, J. (1993). Lag sequential analysis as a tool for functional analysis of student disruptive behavior in classrooms. *Journal of Emotional and Behavioral Disorders, 1*, 138–148.

Gunter, P. L., Shores, R. E., Jack, S. L., Denny, R. K., & DePaepe, P. A. (1994). A case study of the effects of altering instructional interactions on the disruptive behavior of a child identified with severe behavior disorders. *Education and Treatment of Children, 17*, 435–444.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107–128. doi:10.3102/10769986006002107

Heron, T. E., Welsch, R. G., & Goddard, Y. L. (2003). Applications of tutoring systems in specialized subject areas: An analysis of skills, methodologies, and results. *Remedial and Special Education, 24*, 288–300. doi:10.1177/07419325030240050401

Hogan, D. M., & Tudge, J. R. H. (1999). Implications of Vygotsky's theory for peer learning. In A.M. O'Donnell & A. King (Eds.), *Cognitive Perspectives on Peer Learning* (pp. 39–65). Mahwah, NJ: Lawrence Erlbaum Associates. doi:10.4324/9781410603715

Hogan, S., & Prater, M. A. (1993). The effects of peer tutoring and self-management training on on-task, academic, and disruptive behaviors. *Behavioral Disorders*, *18*, 118–128.

Holecek, D. L. (2012). *Cross age/cross disability peer tutoring: A strategy for math instruction* (Unpublished doctoral dissertation). University of Wisconsin-Superior, Superior, WI.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2), 165–179. doi:10.1177/001440290507100203

Horner, R. H., Sugai, G., Todd, A. W., & Lewis-Palmer, T. (2005). School-wide positive behavior support. In L. Bambara & L. Kern (Eds.), *Individualized supports for students with problem behaviors: Designing positive behavior plans* (pp. 359 – 390). New York: Guilford Press.

Hughes, T. A., & Fredrick, L. D. (2006). Teaching vocabulary with students with learning disabilities using classwide peer tutoring and constant time delay. *Journal of Behavioral Education, 15*(1), 1–23. doi:10.1007/s10864-005-9003-5

Institute of Education Sciences (2014). *What works clearinghouse: Procedures and standards handbook (Version 3.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/handbooks.

Johnson, M., & Bailey, J. S. (1974). Cross-age tutoring: Fifth graders as arithmetic tutors for kindergarten children. *Journal of Applied Behavior Analysis, 7*, 223–232.

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*, 82-88.

Jordan, N. C., Kaplan, D., & Hanich, L. B. (2002). Achievement growth in children with learning difficulties in mathematics: Findings of a two-year longitudinal study. *Journal of Educational Psychology, 94*, 586–597. doi:10.1037/0022-0663.94.3.586

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

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research and Practice, 22*, 36-46.

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

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850–867. doi:10.1037/a0014939

Jun, S. W., Ramirez, G., & Cumming, A. (2010). Tutoring adolescents in literacy: A meta-analysis. *Journal of Education, 45*, 219-238. doi:10.7202/045605ar

Kauffman, J. M. (2001). *Characteristics of emotional and behavioral disorders of children and youth* (7th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Kavkler, M., Aubrey, C., Tancig, S., & Magajna, L. (2000). Getting it right from the Start? The influence of early school entry on later achievements in mathematics. *European Early Childhood Education Research Journal*, *8*(1), 75–93.

Kazdin, A. E. (2011). Single-case research designs: Methods for clinical and applied settings. New York, NY: Oxford University Press.

Kennedy, C. H. (2005). Single-case designs for educational research. Boston, MA: Allyn & Bacon.

Kern, L. (2015). Addressing the needs of students with social, emotional, and behavioral problems: Reflections and visions. *Remedial and Special Education, 36*, 24–27. doi:10.1177/07419232514554104

Kleemans, T., Segers, E., & Verhoeven, L. (2011). Cognitive and linguistic precursors to numeracy in kindergarten: Evidence from first and second language learners. *Learning and Individual Differences, 21*, 555-561.

Kohler, F. W., Strain, P. S., Hoyson, M., and Jamieson, B. (1997). Merging naturalistic teaching and peer-based strategies to address the IEP objectives of preschoolers with autism: An examination of structural and child behavior outcomes. *Focus on Autism and Other Developmental Disabilities, 12*(4), 196–206. doi:10.1177/108835769701200402

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

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*(1), 26–38. doi:10.1177/0741932512452794

Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs. *Remedial and Special Education, 24*, 97-114.

Landrum, T. J., Tankersley, M., & Kauffman, J. M. (2003). What's special about special education for students with emotional or behavioral disorders? *Journal of Special Education, 37*, 148–156. doi:10.1177/00224669030370030401

Lane, P., Pollack, C. and Sher, N. (1972). Remotivation of disruptive adolescents. *Journal of Reading*, *15*, 351–364. Retrieved from http://www.jstor.org.ezproxy.lib.utexas.edu/stable/23890485

Lazerson, D. B. (1980). I must be good if I can teach! peer tutoring with aggressive and withdrawn children. *Journal of Learning Disabilities*, *13*, 152–157. doi:10.1177/002221948001300309

Lazerson, D. B. (2005). Detention home teens as tutors: A cooperative cross-age tutoring pilot project. *Emotional & Behavioural Difficulties*, *10*(1), 7–15. doi:10.1177/1363275205050881

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

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*, 1723–1743. doi:10.1111/j.1467-8624.2007.01087.x

Maher, C. A. (1982). Behavioral effects of using conduct problem adolescents as cross-age tutors. *Psychology in the Schools, 19*, 360–364. doi:10.1002/1520-6807(198207)

Maher, C. A. (1984). Handicapped adolescents as cross-age tutors: Program description and evaluation. *Exceptional Children, 51*(1), 56–63. doi:10.1177/001440298405100108

Malmgren, K., Edgar, E., & Neel, R. S. (1998). Postschool status of youths with behavioral disorders. *Behavioral Disorders*, *23*(4), 257–263.

doi:10.1177/019874299802300403

Mastropieri, M. A., Scruggs, T., Norland, J. J., Berkeley, S., McDuffie, K., Tornquist, E. H., et al. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high-stakes tests. *Journal of Special Education, 40*, 130–137. doi:10.1177/00224669060400030101

Mastropieri, M. A., Scruggs, T. E., Spencer, V., & Fontana, J. (2003). Promoting success in high school world history: Peer tutoring versus guided notes. *Learning Disabilities Research & Practice, 18*, 52–65. doi:10.1111/1540-5826.00057

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*(4), 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x.

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

Miller, D., Topping, K., & Thurston, A. (2010). Peer tutoring in reading: the effects of role and organization on two dimensions of self-esteem. *The British Journal of Educational Psychology*, *80*(3), 417–433. doi:10.1348/000709909X481652

Morgan, P. L., & Farkas, G. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*, 306-321.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Who is at risk for persistent mathematics difficulties in the United States? *Journal of Learning Disabilities, 49*, 305– 319.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42*, 306–321. doi:10.1177/0022219408331037

National Council of Teachers of Mathematics. (2008). *Mathematics education in the United States*. *Reston, VA: Author*.

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

National Assessment of Educational Progress. (2015). *Trend in fourth-grade NAEP mathematics average scores for students who are identified as students with disabilities*. Retrieved from http://www.nationsreportcard.gov/reading_math_2015/#mathematics/groups?grade=4

Okilwa, N. S. A., & Shelby, L. (2010). The effects of peer tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the literature. *Remedial and Special Education, 31*(6), 450–463. doi:10.1177/0741932509355991

Oliver, R., & Reschly, D. (2010). Special education teacher preparation in classroom management: Implications for students with emotional and behavioral disorders. *Behavioral Disorders, 35*(3), 188-199. Retrieved from http://www.jstor.org.ezproxy.lib.utexas.edu/stable/43153818

Ollendick, T. H., Weist M. D., Borden, M.C., & Greene, R. W. (1992). Sociometric status and academic, behavioral, and psychological adjustment: A five-year longitudinal study. *Journal of Consulting and Clinical Psychology, 60*(1):80–87

Parsonson, B., & Baer, D. (1978). The analysis and presentation of graphic data. In T. Kratochwill (Ed.), *Single subject research: Strategies for evaluating change* (pp. 105–165). New York, NY: Academic Press.

Pasnak, R., Kidd, J. K., Gadzichowski, M. K., Gallington, D. A., Saracina, R. P., & Addison, K. T. (2009). Promoting early abstraction to promote early literacy and numeracy. *Journal of Applied Developmental Psychology, 30*, 239-249.

Penno, D. A., Frank, A. R., & Wacker, D. P. (2000). Instructional accommodations for adolescent students with severe emotional or behavioral disorders. *Behavioral Disorders, 25*(4), 325–343.

Piaget, J. (1977). *Sociological studies*. London: Routledge.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*, 375–394. doi: 10.1111/j.1467-8624.2007.01131.x

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology, 32*, 146 –159. doi:10.1016/j.appdev.2011.02.005

Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology, 104*, 661–672.

Reijntjes, A., Kamphuis, J. H., Prinzie, P., Boelen, P. A., van der Schoot, M., & Telch, M. J. (2011). Prospective linkages between peer victimization and externalizing problems in children: A meta-analysis. *Aggressive Behavior, 37*, 215–222. doi:10.1002/ab.20374

Reijntjes, A., Kamphuis, J. H., Prinzie, P., & Telch, M. J. (2010). Peer victimization and internalizing problems in children: A meta-analysis of longitudinal studies. *Child Abuse & Neglect, 34*, 244–252. doi:10.1016/j.chiabu.2009.07.009

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review*, *17*(4), 327–362. doi:10.1007/s10648-005-8137-2

Roeders, P. (1995). Learning together – Reflections on the basics of effective small-group-learning. In R. Olechowski, & G. Khan-Svik (Eds.), *Experimental research on teaching and learning* (pp. 309–321). Frankfurt a. M, Germany: Lang.

Scammacca, N., Roberts, G., Vaughn. S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). *Interventions for adolescents struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schleppegrell, M. J. (2010). Language in mathematics teaching and learning. In J. N. Moschkovich (Ed.), *Language and mathematics education: Multiple perspectives and directions for research* (pp. 73-112).

Schunk, D. H. (1998). Peer modeling. In Topping, K., and Ehly, S. (eds.), *Peer-Assisted Learning* (pp. 185-202). Erlbaum, Mahwah, NJ.

Scruggs, T. E., & Osguthorpe, R. T. (1986). Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. *Psychology in the Schools, 23*, 187–193. doi:10.1002/1520-6807

Siegler, R. S., & Booth, J. (2004). Development of numerical estimation in young children. *Child Development,* 75, 428 – 444. doi:10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation: A review. *Psychological Science, 14*, 237-250.

Siegler, R. S., & Ramani, G. B. (2008). Playing board games promotes low-income children's numerical development. *Developmental Science, 11*, 655– 661. doi:10.1111/j.1467-7687.2008.00714.x

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

Smead, V. S. (1984). Self-fulfilling prophecies in the classroom: Dead end or promising beginning? *Alberta Journal of Educational Research, 30*, 145–156.

Soulis, S., & Floridis, T. (2010). Sources of stress for Greek students with intellectual disabilities attending mainstream schools. *Support for Learning, 25*, 74–80. doi:10.1111/j.1467.2010.01443.x

Spencer, V. G. (2006). Peer tutoring and students with emotional or behavioral disorders: A review of the literature. *Behavioral Disorders*, *31*(2), 204–222.

Spencer, V. G., Simpson, C. G., & Oatis, T. L. (2009). An update on the use of peer tutoring and students with emotional and behavioural disorders. *Exceptionality Education International*, *19*(1), 2–13.

Stock, P., Desoete, A., & Roeyers, H. (2009). Detecting children with arithmetic disabilities from kindergarten: Evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities, 43*, 250-268.

Stowitschek, C. E., Hecimovic, A., Stowitschek, J. J., & Shores, R. E. (1982). Behaviorally disordered adolescents as peer tutors: Immediate and generative effects on instructional performance and spelling achievement. *Behavioral Disorders, 7*, 136–148.

Strain, P. S., & Kohler, F., (1999). Peer-mediated interventions for young children with autism: A 20-year retrospective. In Ghezzi, P. M., Williams, W. L., & Carr, J. E. (Eds.) *Autism: Behavior analytic perspectives* (pp. 189-211). Reno, NV.: Context Press.

Sugai, G., & Horner, R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245–259.

Thomas, E. J., and Biddle, B. J. (1966). The nature and history of role theory. In Biddle, B. J., and Thomas, E. J. (eds.), *Role Theory: Concepts and Research,* Wiley, New York, pp. 3–20.

Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of*

Learning Disabilities, 44, 521-532.

Top, B. L., & Osguthorpe, R. T. (1987). Reverse-role tutoring: The effects of handicapped students tutoring regular class students. *The Elementary School Journal*, *87*(4), 413–423. doi:10.1086/461505

Topping, K. (1996). Effective peer tutoring in further and higher education: A typology and review of the literature, *Higher Education, 32*, 321–345. doi:10.1007/bf00138870

Topping, K., & Ehly, S. (1998). *Peer-assisted learning*. Mahwah, NJ: Lawrence Erlbaum. doi:10.4324/9781410603678

Trout, A. L., Nordness, P. D., Pierce, C. D., & Epstein, M. H. (2003). Research on the academic status of children with emotional and behavioral disorders: A review of the literature from 1961 to 2000. *Journal of Emotional and Behavioral Disorders, 11*(4), 198–210. doi:10.1177/10634266030110040201

Turner, R. H. (2002). Role theory. In Turner, J. H. (ed.), *Handbook of Sociological Theory* (pp. 233–254). Kluwer, New York.

University of Texas System/Texas Education Agency. (2009). *Texas Early Mathematics Inventory–AIM Checks*. Austin, TX: University of Texas System/Texas Education Agency.

Utley, C. A., & Mortweet, S. L. (1997). Peer-mediated instruction and interventions. *Focus on Exceptional Children, 29*(5), 1–23. doi:10.1177/074193250102200101

Van Luit, J. E. H., & Schopman, E. A. M. (2000). Improving early numeracy of young children with special educational needs. *Remedial and Special Education, 21*, 27-40.

Vaughn, S., Klingner, J. K., & Bryant, D. P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Remedial and Special Education*, *22*(2), 66–74.

doi:10.1177/074193250102200201

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wagner, J. A. (1995). Studies of individualism-collectivism: Effects on cooperation in groups. *The Academy of Management Journal*, *38*(1), 152–172. doi:10.2307/256731

Walker, H. M. (1995). *The acting-out child: Coping with classroom disruption* (2nd ed.). Longmont, CO: Sopris West.

Watts, G. W., & Bryant, D. P. (2017). *Promoting number sense through number line board games delivered by older students with EBD* (Unpublished pilot study).

Watts, G. W., Bryant, D. P., & Carroll, M. L. (2018). Students with EBD as cross-age tutors: A synthesis of the literature. *Behavioral Disorders, XX*.

Wehby, J. H., Symons, F. J., Canale, J. A., & Go, F. J. (1998). Teaching practices in classrooms for students with emotional and behavioral disorders: Discrepancies between recommendations and observations. *Behavioral Disorders, 24*(1), 51-56.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology, 44*, 588 –596. doi:10.1037/0012-1649.44.2.588

Wood, S. J., & Cronin, M. E. (1999). Students with emotional/behavioral disorders and transition planning: What the follow-up studies tell us. *Psychology in the Schools, 36*, 327–345.

Yoon, J. S., Barton, E., & Taiariol, J. (2004). Relational aggression in middle school: Educational implications of developmental research. *The Journal of Early Adolescence*, *24*(3), 303–318. doi:10.1177/0272431604265681

Zhou, X., & Wang, B. (2004). Preschool children’s representation and understanding of written number symbols. *Early Child Development and Care, 174*, 253-266.

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*, 73–101.