Mississippi State University
Scholars Junction

Theses and Dissertations Theses and Dissertations

12-13-2008

Responsiveness of elementary-aged students, with and without specific learning disabilities, to interventions for mathematics calculation

Masanori Ota

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
Ota, Masanori, "Responsiveness of elementary-aged students, with and without specific learning disabilities, to interventions for mathematics calculation" (2008). *Theses and Dissertations*. 3810.
https://scholarsjunction.msstate.edu/td/3810

This Dissertation - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

RESPONSIVENESS OF ELEMENTARY-AGED STUDENTS, WITH AND WITHOUT SPECIFIC LEARNING DISABILITIES, TO INTERVENTIONS FOR MATHEMATICS CALCULATION

By

Masanori Ota

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Educational Psychology
in the Department of Counseling and Educational Psychology

Mississippi State, Mississippi

December 2008

Copyright by

Masanori Ota

2008

RESPONSIVENESS OF ELEMENTARY-AGED STUDENTS, WITH AND WITHOUT SPECIFIC LEARNING DISABILITIES, TO INTERVENTIONS FOR MATHEMATICS CALCULATION

By

Masanori Ota

Approved:

Carlen Henington Associate Professor of Educational Psychology (Director of Dissertation)	Richard A. Doggett Associate Professor of Educational Psychology (Committee Member)
Harrison D. Kane Assistant Professor of Educational Psychology (Committee Member)	Joe Ray Underwood Professor Emeritus of Counselor Education (Committee Member)
Lynne S. Arnault Professor of Special Education (Committee Member)	Glen R. Hendren Professor and Graduate Coordinator of Counseling and Educational Psychology

Richard Blackbourn
Dean of the College of Education

Name: Masanori Ota

Date of Degree: December 12, 2008

Institution: Mississippi State University

Major Field: School Psychology

Major Professor: Dr. Carlen Henington

Title of Study: RESPONSIVENESS OF ELEMENTARY-AGED STUDENTS, WITH AND WITHOUT SPECIFIC LEARNING DISABILITIES, TO INTERVENTIONS FOR MATHEMATICS CALCULATION

Pages in Study: 217

Candidate for Degree of Doctor of Philosophy

The Response to Intervention (RtI) model is an identification model for Specific Learning Disability (SLD), one of the 13 disability categories identified under the Individual with Disabilities Education Act (IDEA) of 2004. The RtI model has been proposed as an alternative model to the discrepancy model (e.g., intelligence quotient-achievement discrepancy model). In the RtI model, students' responsiveness (e.g., levels of performance and slopes of progress) yields their eligibility for special education. However, to date, research that examined the validity of the RtI model (e.g., examination of intervention responsiveness with students with academic deficits) has been limited in the area of mathematics.

The purpose of this study was to examine the responsiveness of elementary-aged students, with and without SLD, to interventions for mathematics calculation. It was hypothesized that students with mathematics deficits would demonstrate progress after receiving an empirically-derived intervention, regardless of their placement in general or special education. It was also hypothesized that students with mathematics deficits would demonstrate satisfaction with intervention procedures and self-efficacy with their progress after receiving an empirically-derived intervention. Students with and without SLD were selected based on specific criteria for this study (e.g., a skill deficit). To examine these hypotheses, for each student, an intervention was selected using an experimental analysis. The effects of the intervention on mathematics calculation were examined using single subject design. Maintenance on instructional materials and generalization from instructional-level to grade-level materials were examined. Social validity (e.g., satisfaction) of interventions and self-efficacy of students were also assessed. The results of the study indicate that empirically-derived interventions were effective in enhancing the calculation skills of students with and without SLD and maintaining their skills during and after the intervention phase. However, the students with and without SLD did not generalize their calculation skills to grade-level materials. The students demonstrated high levels of satisfaction with the interventions at the end of the interventions and enhanced their self-efficacy across the study. The study partially supported the validity of the RtI model in the area of mathematics such that the RtI model may be reliable in identification of students with SLD in mathematics calculation.

ACKNOWLEDGEMENTS

I thank Dr. Carlen Henington for her continuous, strong support for my dissertation. Discussion with Dr. Henington was always insightful and helpful for me to develop and conduct my research (e.g., research rationale, hypotheses, design, discussion, etc.). I thank my committee members (Drs. Tony Doggett, Harrison Kane, Joe Ray Underwood, and Lynne Arnault) and Dr. Kristin Johnson-Gros for their support for my dissertation. Their suggestions were also helpful for me to conduct this study in a reliable manner. I thank Melissa Harber, Holly Paczac, and Rebecca Roberts for their time assisting with my dissertation (e.g., reliability assessment for intervention integrity and scoring). Their contribution was crucial to the completion of this study. I thank the school personnel who helped me conduct this study in school settings (e.g., the special education coordinator who helped me find schools of potential participants, principals who gave me permission to conduct the study in their schools, teachers who nominated students and sent a parental consent form to each student's parent, etc.) and the parents of the students who gave me permission to work with their child. Finally, I thank the students, who participated in this study, for their efforts to complete mathematics worksheets in each session. Their efforts always made me motivated to continue with this study. It was also enjoyable for me to meet with them each time. I hope all of these students will continue to demonstrate progress in their schools.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

LITERATURE REVIEW

The purpose of this study was to examine the responsiveness of elementary-aged students, with and without specific learning disability, to interventions for mathematics calculations. In this section, previous research on the following issues is reviewed: (a) specific learning disabilities, (b) identification of specific learning disabilities, (c) curriculum-based measurement, (d) academic and mathematics intervention, (e) brief experimental analysis, (f) generalization and maintenance, (g) statement of the problems, (h) purpose of the study, and (i) hypotheses.

Specific Learning Disabilities

Specific Learning Disability in Mathematics Calculation

Mathematics skills such as basic calculations (e.g., addition) and problem solving skills (e.g., mathematics reasoning) are crucial for an individual's success in life. The importance of mathematics education has been emphasized for more than a decade (Carnine, Jones, & Dixon, 1994; Rivera, 1997; Woodward, 2004). However, despite increasing societal demands for strong mathematics ability, some students continue to be academically at-risk in mathematics or have been identified with learning disabilities in mathematics.

Specific learning disability (SLD; see Appendix A) can be identified in one of the 13 categories under the Individuals with Disabilities Education Improvement Act (IDEA) of 2004. In mathematics, two possible learning disabilities may be determined: mathematics calculation (i.e., lack of fluency in basic facts) and mathematics reasoning (i.e., lack of skills to apply calculation skills in solving mathematics word problems or to understand the context of a mathematics problem). Kosc (1974) stated that approximately 6% of children are suspected to have developmental dyscalculia (i.e., a disability in calculation). Geary (2004) stated that approximately 5% to 8% of school-aged children have deficits in cognitive abilities related to mathematics learning disabilities (i.e., SLD in mathematics calculation and mathematics reasoning) such as fact recall, working memory, and problem solving. Kavale and Reece (1992) stated that mathematics was the second highest reason for identification of learning disabilities based on data gathered in record reviews of 917 students with SLD (e.g., reading, mathematics, writing, language, etc.) from preschool/kindergarten through high school in Iowa in 1990. Additionally, Kavale and Reece stated that 14% of lower elementary students (first to third grades) were identified with SLD, whereas 35% were so identified in the upper elementary grades (fourth to sixth grades). This indicated that elementary students became more academically at-risk as they matured.

It is known that as students with unresolved reading difficulties mature, they become increasingly behind in literacy acquisition relative to students who read fluently. This phenomenon, called the Matthew Effect (Stanovich, 1986), might be also found in mathematics and, in the case of mathematics, is due to the fact that a lack of basic calculation skills (e.g., addition and subtraction) hinders acquisition of advanced mathematics skills (e.g., problem solving, skills in higher levels of mathematics such as algebra and geometry). Therefore, it is crucial to identify students with SLD in mathematics calculation early and to implement interventions to specifically address their deficits.

Aspects of SLD in Mathematics Calculation

Researchers have determined that cognitive deficits in mathematics calculation are likely to cause poor calculation performance (e.g., lack of automaticity). In this section, the following aspects of SLD in mathematics calculation are reviewed: (a) cognitive deficits related to mathematics difficulty, (b) lack of automaticity with mathematics facts, (c) error performance and poor strategy use, and (d) comorbidity of deficits in mathematics calculation and reading.

Cognitive deficits related to mathematics difficulty. Cognitive abilities such as working memory and short-term memory affect accurate and fluent calculation. Geary (1993) stated that students with SLD in mathematics calculation have deficits in cognitive skills used in calculation such as counting knowledge (e.g., counting from one such as "one, two," and then "three, four, five" to reach the answer of 2+3; counting from the next number such as "three, four, five" to reach the answer of 2+3), working memory, and counting speed, all of which are essential to mathematics performance. Additionally, deficits in memory processes such as encoding (e.g., an individual's representation of an event; Bjorklund, 2005), storage, and retrieval hinder acquisition and utilization of mathematics skills. For example, students who have memorized mathematics facts retrieve them automatically, whereas students with memory deficits may have difficulty memorizing mathematics facts and rely on finger counting even for simple calculations (e.g., one-digit plus one-digit addition). This reliance leads to slow and/or inaccurate mathematics performance. In fact, memorization of mathematics facts is crucial for fluent calculation (Hasselbring, Goin, & Bransford, 1988; Pellegrino & Goldman, 1987).

Lack of automaticity with mathematics facts. Automaticity indicates fluent and accurate performance without the need to expend excess memory processes. Pellegrino and Goldman (1987) explained this process using the term "expertise" (p. 24) in that shifting from procedural knowledge (knowledge of "how to") to declarative knowledge (knowledge of facts) strengthens the network of mathematics facts, yielding automaticity. Automaticity in calculation is crucial to acquire higher levels of mathematics skills (e.g., advanced mathematics calculation using multiple steps, mathematics reasoning skills). However, researchers have shown that students with SLD in mathematics calculation lack automaticity in mathematics facts (e.g., Gersten, Jordan, & Flojo, 2005; Pellegrino & Goldman, 1987), and that building automaticity in mathematics facts is crucial to promote mathematics performance (Hasselbring et al., 1988). Additionally, repeated practice with mathematics facts has been shown to be effective in developing automaticity in academically at-risk students (e.g., Cates, 2005; Henington et al., 2006).

Error performance and inferior strategy use. Students with SLD in mathematics calculation are likely to make errors and use inferior strategies (e.g., finger counting) due to their lack of automaticity in mathematics facts. For example, Russell and Ginsburg (1984) compared the mathematics skills (e.g., calculation skills) of three groups of students: (a) fourth grade students with mathematics disabilities (MD); (b) fourth grade students without MD; and (c) third grade students without MD. Russell and Ginsburg determined that the fourth grade MD group had skills such as number sense and base ten concepts, but had deficits in skills such as counting large numbers and addition fact retrieval. Additionally, compared to the two groups without MD, the students in the fourth grade MD group used wrong operations (e.g., addition instead of subtraction) and made simple miscalculations (e.g., addition errors). Russell and Ginsberg concluded that "unusual difficulty with the simplest number facts" (p. 241) is the main aspect of MD. This study indicates that error performance and inferior strategy use are aspects of students with SLD in mathematics calculation.

Comorbidity of SLD in mathematics and reading. Reading difficulties have been shown to affect the mathematics performance of students diagnosed with SLD in mathematics (e.g., calculation and/or reasoning). Jordan and colleagues (Jordan & Hanich, 2000, 2003; Jordan, Hanich, & Kaplan, 2003) have demonstrated that comorbidity of SLD in mathematics and reading hinders students' calculation. For example, Jordan and Hanich (2000) compared the mathematics performance of four groups of students: (a) typical students, (b) students diagnosed with SLD in reading, (c) students diagnosed with SLD in mathematics only, and (d) students diagnosed with SLD comorbidity (i.e., SLD in both mathematics and reading). The researchers found that both the SLD in mathematics and SLD comorbidity groups performed worse in calculation and story problems and used more primitive strategies (e.g., finger counting) in mathematics facts than the typical group. However, the comorbidity group used these primitive strategies less accurately (e.g., inaccurate finger counting) and made more errors in mathematics facts than did the SLD in mathematics only group. Jordan and Hanich concluded that learning difficulties in the comorbidity group were more pervasive in mathematics (e.g., deficits in both mathematics facts and mathematics reasoning) than those in mathematics only group and that the students' reading deficits may have hindered their understanding of mathematics concepts and use of effective strategies.

Overall, reading difficulties of students with co-occurring learning disabilities in reading and mathematics may affect their performance accuracy (e.g., frequent errors) and/or use of strategies (e.g., primitive counting strategies) in mathematics. Therefore, for these students, mathematics interventions should address mathematics deficits that are likely to be affected by their reading deficits (e.g., verbal counting).

Identification of SLD in Mathematics Calculation

A recent paradigm shift has occurred in the identification process of students with SLD in efforts to meet their educational needs. Although currently, the discrepancy model (e.g., intelligence quotient (IQ)-achievement, achievement-achievement), in which a student is determined to meet the criteria for SLD when the student's achievement score on one or more areas is significantly below his or her IQ score or achievement scores in other areas (e.g., one standard deviation), is used in many educational settings, the discrepancy model is not likely to lead to effective interventions for these students (Shaprio, 1996, 2004). Rather, researchers (e.g., Brown-Chidsey & Steege, 2005; Fuchs

& Fuchs, 1998; Gresham, 2002) have recommended an alternative SLD identification model, the Response-to-Intervention (RtI) model, in which only those students who do not appropriately respond to academic interventions with a certain period of time (i.e., those identified as nonresponders; Fuchs & Fuchs, 2002) are identified with SLD. Additionally, educational laws, such as No Child Left Behind (NCLB) of 2001 and the Individuals with Disabilities Act (IDEA) of 2004, have included RtI as a potential indicator in SLD identification (Brown-Chidsey & Steege, 2005; Fletcher, Coulter, Reschly, & Vaughn, 2004). Researchers have stated that the RtI model has treatment validity, as opposed to traditional discrepancy model, in that interventions are implemented with students who are academically at-risk and their responsiveness rather than discrepancy (IQ-achievement or achievement-achievement discrepancy) yields special education eligibility (Fuchs & Fuchs, 1998). Implementation of preventive, empirically-based interventions in general education has been also emphasized in the field of school psychology (Kratochwill & Shernoff, 2003). In the area of reading, some researchers have also demonstrated that academically at-risk students who showed a discrepancy in standardized test scores (e.g., potential students with SLD) responded to an intervention (e.g., Vellutino et al., 1996). However, research in the intervention responsiveness of students with SLD in mathematics calculation to instruction has been limited (e.g., Fuchs, Fuchs, & Prentice, 2004).

In summary, in the RtI Model, intervention responsiveness, rather than a discrepancy between standardized test scores (e.g., IQ-achievement discrepancy), yields SLD identification. The assessment of intervention responsiveness in the RtI model has also been demonstrated to be a valid methodology in identification of students with SLD (i.e., the validity of the RtI model) predominantly in the area of reading (e.g., Vellutino et al., 1996). Therefore, researchers should also examine the validity of the RtI model for SLD in mathematics calculation.

Summary for SLD in Mathematics Calculation

Students with SLD in mathematics calculation commonly have deficits in cognitive skills (e.g., automaticity). IQ is a measure of an individual's intellectual ability composed of specific cognitive abilities, including verbal, spatial, speed of processing, and memory (Bjorklund, 2005). Additionally, an individual's IQ is assessed using an IQ test and IQ scores are considered to predict the individual's current and future achievement. For example, in mathematics, memory and speed of processing may be associated with memorization of facts and fluency in basic calculation (e.g., addition and subtraction). A verbal ability may also affect problem solving in mathematics (e.g., mathematics reasoning problems such as word problems). Students' current levels of academic skills are also assessed using achievement tests in educational settings. However, a student's intelligence or achievement test scores, or a discrepancy in these test scores between the student's strength and weakness areas, may not predict his or her future progress including intervention responsiveness (e.g., Fuchs et al., 2004; Vellutino et al., 1996). In the past, students with difficulty in mathematics facts are likely to be identified with SLD in mathematics calculation based on the discrepancy model. However, this assessment model is unlikely to lead to sufficient information to assist in the development of a specific intervention to remediate the student's deficits (Gresham &

Witt, 1997). Therefore, further research regarding valid alternative identification models in SLD in mathematics calculation that maximize the potentials of effective interventions (e.g., RtI model) is warranted.

In the following sections, two identification models for SLD (discrepancy model and RtI model) and studies regarding these two models are reviewed.

Identification of Specific Learning Disabilities

Identification models of specific learning disabilities have been controversial. To date there are two main SLD identification models: (a) discrepancy model and (b) RtI model. Each of the identification models and related studies are reviewed in this section.

Discrepancy Model

The discrepancy model is a SLD identification methodology, in which a student whose achievement score in one area is a certain level below his or her IQ score (i.e., IQ-achievement discrepancy) or achievement score(s) in other area(s) (i.e., achievement-achievement discrepancy) is identified with SLD. The discrepancy model has been used in SLD identification in most of the states (Mercer, Jordan, Allsopp, & Mercer, 1996) and one standard deviation discrepancy is commonly used; for example, a discrepancy of 15 points between IQ and achievement scores is used in the State of Mississippi (Mississippi Department of Education, Office of Special Education, 2002).

However, the validity of the discrepancy model has been questioned (Fuchs et al., 2004; Siegel, 1989, 1992; Vellutino et al., 1996; Ysseldyke, Algozzine, Shinn, & McGue, 1982). For example, in the IQ-achievement discrepancy model, an IQ score may not accurately predict students' performance (e.g., reading; Siegel, 1989, 1992). Furthermore, an IQ score neither differentiates low achievers nor students with SLD (Ysseldyke et al., 1982). Generally, in the discrepancy model, standardized test results (e.g., IQ and achievement test scores or discrepancy of these scores) may not yield interventions for students with SLD (Gresham & Witt, 1997). Many researchers (e.g., Fletcher et al., 2002) believe that testing results from a single testing opportunity may not yield valid decisions (e.g., special education eligibility). In previous research, students identified with SLD based on standardized test scores demonstrated progress when receiving interventions (e.g., in reading, Vellutino et al.). These facts lead to questions about the implementation of the discrepancy model for SLD identification in school.

As reviewed, inadequacy of the discrepancy model proposes an essential question regarding SLD identification: What is the definition of SLD? To answer this question, researchers have proposed an alternative identification model, the RtI model, in which students' responsiveness to intervention is examined for SLD identification. The RtI model and its validity are reviewed in the following section.

RtI Model

The RtI model is defined as "a systematic and data-based method for identifying, defining, and resolving students' academic and/or behavior difficulties" (Brown-Chidsey & Steege, 2005, p. 2). In the RtI model, only students who do not respond to interventions (i.e., nonresponders) are identified with SLD. Fuchs and Fuchs (1998) proposed a "dual discrepancy" (p. 205) model, one version of the RtI model in which students who demonstrate both low levels of performance and low rates of progress relative to peers are identified with SLD. Thus, this model states that students who respond to interventions should be taught in general education, despite their low achievement. These students may also be considered to be slow learners as stated in No Child Left Behind (2001). Cooter and Cooter (2004) stated that slow learners are those who demonstrate low IQ and low achievement scores (e.g., IQ scores ranging from 70 to 85). These students may be also in a "failure to thrive" (Fuchs & Fuchs, 2002, p. 79) situation such that intense interventions that are implemented for a certain period of time are needed for them to demonstrate progress. As such, implementation of interventions within the RtI model may best address the educational needs of these students, before their special education eligibility is considered.

Researchers have also stated the three-tier (e.g., Brown-Chidsey & Steege, 2005) or four-tier (e.g., Gresham, 2002) RtI model, in which interventions are implemented with academically at-risk students such that their responsiveness is evaluated in each tier. Additionally, in both models, more intense interventions (e.g., small group or individual interventions) are implemented with academic at-risk students in later tiers (e.g., the second, third, and/or fourth tiers). In either model, only students who failed to respond to interventions in the last tier may be considered for a diagnosis of SLD.

In summary, the RtI model is a preventive model for SLD in academics (Fuchs, 2005; Fuchs & Fuchs, 2001). The RtI model also meets the call for evidence-based practice in school psychology (e.g., Kratochwill & Shernoff, 2003) and is a potential model for SLD identification.

Two Problem-Solving Frameworks for the RtI Model

Within the RtI model, there is an emphasis on problem-solving and intervention, rather than disability diagnoses for students who struggle academically. To this end, researchers have proposed problem-solving frameworks for the RtI model. In each of the frameworks, intervention intensity increases as levels proceed and students' responsiveness to intervention is evaluated at each of the levels.

Three-tier model. Brown-Chidsey and Steege (2005) proposed the three tiers of intervention for the RtI model: (a) general instruction and assessment for all students (Tier 1) (e.g., classwide curriculum-based assessment), (b) supplementary instruction (e.g., small group intervention) and assessment for some academically at-risk students who do not respond to general instructions in Tier 1 (i.e., Tier 2) (e.g., those in intervention and monitoring), and (c) specialized instruction and monitoring for academically at-risk students who do not respond to supplementary instructions in Tier 2 (i.e., Tier 3) (e.g., those in comprehensive assessment). In the framework, all assessment and interventions in the three tiers are conducted in general education and only students who failed to respond to intervention at Tier 3 are potentially eligible for special education.

Four-tier model. Researchers have also presented a four-level problem-solving framework, the Heartland Problem Solving Approach, conducted in Iowa (e.g., Gresham, 2002). The four levels are (a) Level I: Primary prevention (class/school-wide interventions), (b) Level II: Secondary prevention (parent/teacher consultation), (c) Level

III: Tertiary prevention (small group/individual interventions), and (d) Level IV: Special education and IEP determination (intense academic remediation). In the framework, an instructional consultation is provided to teachers, and only students who failed to respond to interventions are referred to the Level IV and decision is considered (e.g., special education eligibility).

In either framework, the RtI model has treatment validity, given that intervention results (e.g., responsiveness) yield educational decisions (Fuchs & Fuchs, 1998). The RtI model has been found to effectively decrease the number of students identified with SLD in Minneapolis (Marston, Muyskens, Lau, & Canter, 2003). The RtI model has been found to be valid for reading (e.g., Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al., 1996) and mathematics (Fuchs et al., 2004). In summary, researchers have stated that the RtI model is a preventive problem-solving framework. The RtI model generally consists of tiers (e.g., three or four tiers), in which intensity of interventions is enhanced as a tire proceeds and only those who do not respond to interventions in the last tier are identified with SLD. The validity of the RtI model has been demonstrated in basic subject areas including reading and mathematics; however, further research is warranted to address this issue in mathematics. Research regarding the validity of the RtI model in mathematics is reviewed in the following sections.

Research on Intervention Responsiveness in the Area of Mathematics

Researchers have demonstrated the validity of the RtI model in SLD identification in mathematics. Specifically, students with SLD in mathematics demonstrate progress when teacher instructions are modified (Fuchs, Fuchs, Hamlett, Phillips, & Bentz, 1994;

Fuchs, Fuchs, Hamlett, & Stecker, 1990, 1991) or when mathematics interventions are implemented with students (Fuchs et al., 2004).

Teacher modification of instruction. Instructional feedback and skill analysis enhance teachers' instruction and students' progress. Fuchs et al. (1990) examined the effects of teachers' use of curriculum-based measurement (CBM; Deno, 1985) with graphed data feedback for and skill analysis of a student's mathematics performance across 15 weeks. Thirty special education teachers were randomly assigned to three groups: (a) CBM with graph feedback and skills analysis, (b) CBM with graph feedback only, and (c) control group (no feedback or analysis). Each experimental or control teacher selected a few students with disabilities (e.g., SLD in mathematics) in his or her class. Additionally, teachers in the graph feedback and skills analysis group evaluated students' progress, targeted skills, and instructional modification on a computer, whereas teachers in the graph feedback only evaluated only student progress on a computer. Control teachers did no feedback or analysis. Results indicate that students in the graph feedback and skills analysis group showed the highest progress, indicating that students with SLD can demonstrate progress if teachers' instructions target their skills. The results of this study indicated that individual instructional modification may enhance the academic performance of students with SLD in mathematics. In fact, individual instructional modification is crucial to enhance students' responsiveness to intervention in the RtI model.

Mathematics interventions with students. Interventions with students also enhance their mathematics performance. Fuchs et al. (2004) examined the response of the 301 third grade students to a 16-week mathematics intervention. The students were classified into a control or experimental group and were further classified into four groups: (a) at risk for both mathematics and reading disabilities (MDR/RDR), (b) at risk for mathematics disability only (MDR-only), (c) at risk for reading disability only (RDR-only), and (d) not at risk (NDR). The intervention consisted of transfer (generalization) instruction and self-regulation. In the transfer instruction, teachers taught their students skill application (e.g., application of problem-solving skills to similar problems), whereas in the self-regulation, students used self-monitoring (e.g., scoring answers using a key) and goal setting procedures. Pretest and posttest were administered before and after the intervention on each performance dimension (e.g., calculation). The results indicated that for experimental group comparison, the three disability groups (e.g., MDR/RDR) improved less than the NDR group on calculation. For experimental and control group comparison, each experimental disability group demonstrated greater improvement than the counter control disability group. This study indicates that students with SLD in mathematics respond to interventions. Additionally, this study differs from previous studies (e.g., Jordan & Hanich, 2000) in that Fuchs et al. (2004) examined students' *responsiveness* rather than performance levels without implementing interventions as in the previous studies, indicating that responsiveness to intervention is a viable measure to identify SLD in mathematics calculation.

Limitations in the Previous Studies in the RtI Model in Mathematics Calculation

Despite the results of previous studies in the RtI model in mathematics in calculation, the results of the previous studies reviewed in this section should be viewed with caution. First, group design was used in most of the studies (e.g., Fuchs et al., 1994; Fuchs et al., 1990, 1991; Fuchs et al., 2004). Rather, single subject design may be used to examine the individual responsiveness of academically at-risk students. Second, students' calculation skills were not assessed using a sensitive measure such as CBM. Third, mathematics calculation was not solely targeted. In Fuchs et al. (2004), mathematics calculations were embedded in mathematics reasoning problems and thus it was not clear whether responsiveness was due to calculation or mathematics reasoning skill. Fourth, interventions were not selected experimentally for each student. Interventions based on experimental analysis may maximize responsiveness (e.g., Duhon et al., 2004). Fifth, social validity (Wolf, 1978) of interventions was not assessed. Interventions with high social validity may enhance students' responsiveness. Finally, few researchers have examined responsiveness of students in both general and special education (e.g., Fuchs et al., 2004). Therefore, further research is warranted to address these issues. Research that addresses these issues may clarify the idiosyncratic responsiveness of each student as well as potential variables that predict students' responsiveness under certain interventions (e.g., Codding et al. 2007; Rhymer et al., 1998).

Summary for Identification of Specific Learning Disabilities

The RtI model has been proposed as an alternative to the discrepancy model (e.g., IQ-achievement or achievement-achievement). To date, the discrepancy model has been used widely in the identification of students with SLD. However, the validity of the discrepancy model has been questioned (Fuchs et al., 2004; Siegel, 1989, 1992; Vellutino et al., 1996; Ysseldyke et al., 1982). For example, assessment results obtained in the discrepancy model may not yield effective interventions that address the deficits of academically at-risk students (i.e., lack of treatment utility) (Gresham & Witt, 1997). Rather, researchers (e.g., Brown-Chidsey & Steege, 2005; Fuchs & Fuchs, 1998; Gresham, 2002) have stated that the RtI model should be used in identification of students with SLD. In the RtI model, interventions are implemented with academically at-risk students in general education, a student's responsiveness is evaluated, and only students who fail to respond to interventions in the final tier are referred for evaluation for SLD (Brown-Chidsey & Steege; Fuchs & Fuchs; Gresham). Specifically, the researchers have stated a three-tier (e.g., Brown-Chidsey & Steege) or four-tier (e.g., Gresham) model, in which more intense interventions are implemented in higher tiers and students' responsiveness is evaluated in each tier. Researchers have demonstrated that students with SLD in mathematics respond to targeted interventions (e.g., Fuchs et al., 1994; Fuchs et al., 1990, 1991; Fuchs et al., 2004). However, further research is warranted to verify the validity and utility of the RtI model in identification of SLD in mathematics and efforts should be made to assist interventionists in selecting specific interventions based on students' individual characteristics. To do so, for example, researchers may examine a student's responsiveness using a single subject design in which an idiosyncratic intervention (e.g., an intervention identified as effective experimentally for each student) is implemented for a certain period and his or her responsiveness is analyzed (e.g., an analysis of level, trend, and variability in data).

For the RtI model to be utilized in schools, it should be researched regarding (a) potential academic interventions, (b) sensitive measures to assess responsiveness, and (c) criteria against which students are identified as nonresponders. CBM, a sensitive measure to responsiveness, and its use in the RtI model are reviewed in the following section.

Curriculum-Based Measurement

CBM General

CBM (Deno, 1985) is a set of measurements used to assess students' academic performance based on their curriculum in basic subject areas (e.g., mathematics, reading, etc.). CBM is a global measure targeting specific skills (e.g., addition in mathematics), is expressed as a rate such as digits correct per minute (DCPM) in mathematics, is sensitive to progress monitoring, and has a number of characteristics making it a valuable tool for educators (e.g., progress monitoring, program evaluation, etc.) (Deno, 1986; Shapiro, 1996, 2004; Shinn, 1989). Given that content levels are always set at or close to students' instructional levels, CBM has no ceiling or floor effects. CBM is advantageous to traditional standardized tests (e.g., intelligence and achievement tests) in that CBM can be repeatedly used; students' progress is evaluated using single subject design methodology; and CBM has treatment utility (i.e., CBM data yield intervention development) (Hayes, Nelson, & Jarrett, 1987). CBM has been demonstrated to be a reliable and valid measure in basic subject areas such as mathematics and reading (Shinn, 1989). For example, in mathematics, CBM has high reliability (Fuchs et al., 1988, cited in Marston, 1989) and construct validity (Shinn & Marston, 1985; Thurber, Shinn, & Smolkowski, 2002).

In mathematics, researchers have also demonstrated that CBM procedures enhance students' achievement and teachers' instructions (Fuchs et al., 1994; Fuchs et al., 1990, 1991). Additionally, CBM can be used for evaluation of students with SLD who are mainstreamed into general education (Fuchs, Fuchs, & Fernstrom, 1993). As such, CBM is a sound measurement tool to aid in examining the responsiveness of students in general and special educations.

CBM in Mathematics and the RtI Model

CBM is a sensitive measurement tool to assess students' response to intervention and, ultimately, identify students with SLD within the RtI model (Brown-Chidsey & Steege, 2005; Fuchs & Fuchs, 1997; Fuchs & Fuchs, 1998). Specifically, growth as determined by increasing trends and/or changes in levels (Hayes, Barlow, & Nelson-Gray, 1999) yield information about a student's responsiveness to intervention. Potential criteria for responsiveness are as follows.

One way to evaluate a student's responsiveness to an intervention across time is to evaluate the slope of the trendline or "the student's actual rate of progress" (Shinn & Bamonto, 1998, p.9). The slope indicates rates of learning over time. For example, in the RtI model, an interventionist may implement an academic intervention with an academically at-risk student for several weeks (e.g., 8 weeks) and calculate a weekly slope as a datum of the student's progress. Fuchs, Fuchs, Hamlett, Walz, and Germann (1993) provided weekly mean slopes in mathematics, reading, and spelling from first through sixth grades. Two-year data of approximately one thousand students from first to sixth grade in general education including mainstreamed students with disabilities (e.g., SLD) were collected across five school districts. For example, in mathematics, the weekly mean slopes for third graders were .42 and .30 in Year 1 and 2, respectively. Based on these results, Fuchs, Fuchs, Hamlett, Walz et al. proposed expected weekly CBM slopes in mathematics for each grade (e.g., .50 DCPM for first to third grade; 1.15 to 1.20 DCPM for the fourth and fifth grades; see Table 1.1). Thus, it is possible to compare the rate of progress of a specific student to expected rates by grade.

Table 1.1 Expected Weekly CBM Slopes in Mathematics for First through Fifth Grades Suggested by Fuchs et al. (1993)

Grade	Expected Weekly CBM Slope
Grade 1-3	.50 DCPM
Grade 4-5	1.15 – 1.20 DCPM

Note. CBM = curriculum-based measurement.

Another method of evaluating a student's response to an intervention is to look for a change in the level of performance. Students' levels of performance are identified based on the standardized CBM criterion (Deno & Mirkin, 1977, cited in Shapiro & Lentz, 1986; see Appendix B). Specifically, in CBM, students' levels of performance are categorized into the three levels: (a) frustrational, (b) instructional, and (c) mastery. Instructional level indicates a level in which students' performance falls within an average range at the grade. Frustrational or mastery level indicates a level in which students' performance fall below or above, respectively, an average range at the grade. For example, in mathematics, students' level is at frustrational if their DCPM are below 10 for first through third grades or if their DCPM is below 20 for fourth through sixth grades. Therefore, students who fall within frustrational level at their grades are academically at-risk. Additionally, students are expected to reach mastery level in each skill at each grade when receiving interventions. Furthermore, students may show a distinct and immediate improvement in performance following implementation of an intervention while remaining with the instructional level. Ideally, this jump in performance would then be followed by an increasing slope of the trend line.

Therefore, researchers can use the CBM slope and level as a basis to distinguish responders from nonresponders in the RtI model in mathematics. Furthermore, based on a dual discrepancy model, students who fall within frustrational level at their grade placement *and* whose weekly slopes are lower than the aspirational goals proposed by Fuchs et al. (1993) are considered to be nonresponders and should be provided with further assessment to determine eligibility for special education services (i.e., Tier 3 in

Brown-Chidsey and Steege's three tiers model or the Level IV in the Heartland Problem Solving Approach).

Summary for CBM

In summary, CBM is a global, sensitive measurement tool used to assess students' academic skills based on their curriculum (Shapiro, 1996, 2004; Shinn, 1989). CBM can be used for various educational purposes (e.g., progress monitoring, program evaluation, etc.) (Deno, 1986). CBM has been demonstrated to be a reliable (Fuchs et al., 1988, cited in Marston, 1989) and valid (Shinn & Marston, 1985; Thurber, Shinn, & Smolkowski, 2002) measure. CBM can be also used to assess students' intervention responsiveness within the RtI model (Brown-Chidsey & Steege, 2005; Fuchs & Fuchs, 1997; Fuchs & Fuchs, 1998). To identify students who need special educational services (e.g., students with SLD), researchers and practitioners implement interventions with academically at-risk students and assess their responsiveness using CBM. Specifically, in responsiveness evaluation, changes in levels of performance and/or rates of progress are analyzed across time using single subject methodologies (Hayes et al., 1999; Shinn & Bamonto, 1998). A research-based CBM norm for each grade (e.g., a mean weekly slope of progress; Fuchs, Fuchs, Hamlett, Walz et al., 1993) can be also used to differentiate between responders and nonresponders. Given the treatment utility of CBM (Hayes et al., 1987), indicating that CBM data can be used to develop effective interventions, and the sensitivity of CBM in monitoring students' progress through interventions, CBM should be utilized when assessing students' responsiveness in the RtI model. Potential academic and mathematics interventions that address calculation in the RtI model are reviewed in the following section.

Academic and Mathematics Interventions

Researchers have identified effective interventions for academically at-risk students which may be implemented in the RtI model. Researchers have stated that interventions that address basic academic skills, provide increased opportunities to respond, and enhance learning rates are most likely to promote students' performance (Skinner, 1998; Skinner, Fletcher, & Henington, 1996). Additionally, it has been demonstrated that an academic intervention is most effective when it addresses student's skill or performance deficit (Skinner). Therefore, the following potential mathematics interventions that address these issues are reviewed in this section: (a) repeated practice, (b) folding-in technique, and (c) immediate corrective feedback. The importance of social validity of mathematics interventions and students' self-efficacy are also discussed.

Repeated Practice

Repeated practice (RP), also called drill and practice, increases opportunities to respond by having the student repeatedly practice the same or similar problems. RP has been demonstrated to promote student's fluency in reading (e.g., Lovitt & Hansen, 1976; Samuels, 1997) and in mathematics (e.g., Cates, 2005; Hasselbring et al., 1988; Henington et al., 2006; Maccini & Hughes, 1997). In mathematics, RP enhances the automaticity of students with learning disability (Hasselbring et al.) and is frequently used for these students (Maccini & Hughes).

Cates (2005) compared the effects of peer and computer drills on the acquisition of basic addition facts of 4 fourth through sixth grade students who had mathematics difficulties. The study involved a BCBC design (B as peer drill and C as computer drill). However, a third peer drill phase was added if a student demonstrated progress, which allowed continuous progress (i.e., a BCBCB design). In the peer drill, younger students (2 fourth grade students) and older students (fifth and sixth grade students) were paired. The tutor presented as many flashcards as possible to a tutee for 3 minutes. When a tutee made an error, a tutor made no correction but showed the same flashcard until the tutee responded correctly. The tutor and tutee roles were switched after 3 minutes. On the other hand, in the computer drill, each student individually responded to flashcards projected on a computer. If the student made an error, the same flashcard was displayed until the student responded correctly. The computer program was continued for 3 minutes. In both of the conditions, the dependent variable was the number of correctly answered flashcards in 3 minutes. Results indicated that, computer drill was more effective for older students, whereas peer drill was more effective for younger students. Cates (2005) stated that one possible reason for this difference is that students' response rates matched problem presentation rates in each intervention (e.g., slow flashcard presentation in peer drill for younger students with slow response rates). Cates concluded that presentation pace of antecedent stimuli should be considered in developing fluency. From this fact, it is expected that slow learners (e.g., students with SLD) likely acquire mathematics facts under slow rates of learning (e.g., repeated untimed practice of worksheets).

Another version of RP is completion of the same or similar mathematics worksheets several times. This procedure is frequently used in reading interventions such as repeated reading (Samuels, 1997). In mathematics, completion of the same or similar mathematics worksheet may enhance memorization of mathematics facts and, thus, fluency in solving mathematics problems. Therefore, RP may enhance learning rates of academically at-risk students in mathematics.

Folding-In Technique

Control of task difficulty by changing ratios of instructional and mastery problems is also an effective method to promote students' academic performance. Gickling and Thompson (1985) stated that task difficulty is categorized into the following three levels based on the ratio of known and unknown items for each student: (a) frustrational (fewer than 70% known items), (b) instructional (70-85% of known items and 15-30% unknown items), and (c) independent (more than 90% known items). Gickling and Thompson also stated that for drill (e.g., calculations), students' performance is enhanced the most effectively when 70% known and 30% unknown items were presented. A ratio of known and unknown items with unknown items interspersed into known items is referred to as folding-in technique (FI; Shapiro, 1996, 2004). FI has been found to effectively promote mathematics calculations (Cooke, Guzaukas, Pressley, & Kerr, 1993; Cooke & Reichard, 1996).

Cooke et al. (1993) conducted three experiments to compare the following different versions of FI in spelling, mathematics, and reading with different participants in each experiment. The following two interventions were compared: (a) 30/70% (30%

unknown and 70% known items) condition and (b) 100% condition (100% unknown items). An alternating treatments design (ATD; Hayes, Barlow, & Nelson-Gray, 1999), a single subject design in which either of the interventions was implemented on each session, was conducted. In the mathematics experiment, 10 flashcards of multiplication facts were administered to 3 elementary students (i.e., three unknown facts and seven known facts in the 30/70% and 10 unknown facts in the 100% condition). Specifically, in either condition, a teacher presented flashcards to students for 2 seconds. If the student made an error, the teacher provided the correct answer, had the student repeat it, and repeated this procedure two more times. Next, the student completed three worksheets containing facts learned in the session. The teacher provided praise for correct responses and corrective feedback for incorrect responses. Then, the teacher presented each flashcard again and marked "+" (correct) or "–" (incorrect) at the back of the flashcard. Three consecutive "+" marks indicated mastery. Finally, the student completed a generalization worksheet in which all the facts assigned to the condition in the session were randomized. A maintenance test was administered every other week, in which 10 cards from all the facts mastered to date, five of which were selected from each of the conditions (30/70% and 100% conditions), were presented to the student. Correct response cards were given to the student with praise, whereas incorrect response cards were replaced as unknown. Results indicated that the 30/70% condition was more effective in promoting students' acquisition of facts and retention on the maintenance test than the 100% condition. Additionally, 2 of the 3 students favored the 30/70% condition. It should be noted that the results for spelling and reading were different; students made similar gains under the 30/70% condition and 100% condition in spelling, whereas 2 of the 3 students gained more under the 100% condition for reading. Given that students were different in each subject, response patterns might be idiosyncratic. Therefore, as Cooke et al. (1993) stated, it is necessary to determine an appropriate ratio of known and unknown problems depending on the student's level of acquisition (e.g., 30/70% for low achievers, 50/50% for moderate achievers, etc.).

The effects of interspersing a different ratio of known and unknown items have been also demonstrated in intervention studies in the area of mathematics (e.g., Cates & Skinner, 2000; Logan & Skinner, 1998; Robinson & Skinner, 2002; Skinner, 2002; Wildmon, Skinner, McCurdy, & Sims, 1999; Wildmon, Skinner, Watson, & Garrett, 2004). Skinner and his colleagues demonstrated that interspersal worksheets (i.e., worksheets including 100% of instructional problems plus 20% to 40% of mastery problems) effectively enhanced accuracy and learning rates in school-aged students and that the students favored interspersal worksheets compared to general worksheets including only 100% instructional problems. Effective ratios in interspersal worksheets have been also examined. For example, Hawkins, Skinner, and Oliver (2005) demonstrated that 1:3 interspersal worksheets (i.e., worksheets with a ration of one interspersal problem per three instructional problems) were more effective in enhancing students' accuracy than were 1:1 interspersal worksheets (i.e., worksheets with a ratio of one interspersal problem per one instructional problem). Hawkins et al. explained the results such that a ratio of 1:3 may have provided students with an appropriate rate of reinforcement to pay attention to tasks compared to a ratio of 1:1 which provides equal, but too much attention on both type of problems.

In summary, interventions that target practice of facts including different ratios of known and unknown facts (e.g., FI, interspersal worksheets, etc.) may be effective in promoting the fact acquisition of students with deficits in mathematics calculation. Given the findings of Hawkins et al., a ratio of 30% unknown and 70% known items used in the FI (i.e., a ratio of approximately 1:2.3) may be effective in enhancing acquisition of mathematics facts in academically at-risk students.

Immediate Corrective Feedback

Immediate corrective feedback (ICF) is a method in which the student is told if his or her response is correct or incorrect immediately after each response. ICF, provided by others or by the student to himself/herself, has been found to promote students' correct performance in mathematics (Bennett & Cavanaugh, 1998; Skinner, Shapiro, Turco, Cole, & Brown, 1992) and word acquisition (e.g., Belfiore, Skinner, & Ferkis, 1995).

Bennett and Cavanaugh (1998) examined the effects of immediate self-correction, delayed self-correction, and no correction on the acquisition of multiplication facts of a fourth grade student with learning disabilities. In this study, the effects of immediate self-correction and no-correction were compared in Experiment 1, and one week later the effects of immediate and delayed self-correction were compared in Experiment 2. In each experiment, the teacher presented multiplication flashcards and corrected the student's incorrect responses. After this instruction, the student completed four five-item tests. However, the teacher's procedures on these tests were different among the three conditions. In the no-correction condition, the student did not correct her responses on the tests. In the immediate self-correction condition, the student self-corrected her responses after completing each test. In the delayed self-correction condition, she self-corrected the four tests after completing all of them. The student took a daily test after the four tests. The student took maintenance tests 3 weeks after the termination of the intervention in each experiment. In each of the experiments, accuracy (e.g., the number of problems answered correctly per minute) and error rates on daily tests were assessed. Results indicated that, in Experiment 1, immediate self-correction was more effective than no-correction for the student's accuracy. In Experiment 2, immediate self-correction was more effective than delayed self-correction for the student's accuracy. Additionally, in both of the experiments, immediate self-correction was more effective in reducing the student's repetition of errors than the other two conditions.

This study indicates that immediate correction is more effective in promoting accuracy and preventing errors than delayed or no correction. Prevention of errors is also necessary for academically at-risk students to enhance their mathematics fluency, given that frequent errors may hinder acquisition of mathematics calculation skills. Therefore, it is expected that ICF promotes calculation accuracy in academically at-risk students. Overall, ICF is effective in enhancing students' accurate academic performance and reducing errors. Error performance has been demonstrated to be one of the common aspects of students with deficits in mathematics (e.g., Russell & Ginsburg, 1984). For these students, ICF, rather than delayed or no feedback, may be one of the effective strategies in enhancing accuracy and automaticity in calculation (e.g., Bennett & Cavanaugh, 1998).

In summary, for students with deficits in calculation, skill-based interventions reviewed in this section may provide opportunities to respond, promote acquisition of facts, and enhance their automaticity (Skinner, 1998; Skinner et al., 1996). Specifically, practice (e.g., Cates, 2005) and feedback (e.g., Bennett & Cavanaugh, 1998; Skinner et al., 1992) components included in these interventions are effective in promoting calculation accuracy and fluency in these students. Although academic performance (e.g., calculation performance) should be a primary dependent variable in intervention studies, researchers may also assess related variables including the acceptability of interventions and students' perception of their academic performance (e.g., self-efficacy) when implementing interventions. Research on social validity and self-efficacy is reviewed below.

Social Validity and Self-Efficacy in Intervention Research

Social validity of an intervention and students' self-efficacy are variables that may affect students' intervention responsiveness and willingness to participate in the intervention. In intervention studies, researchers may examine both students' academic skills and related variables including social validity and self-efficacy. In this section, literatures regarding social validity (e.g., various theoretical models of social validity) and self-efficacy (e.g., association between academic performance and self-efficacy) are reviewed.

Social Validity

Social validity refers to the extent to which intervention procedures and effects are acceptable and meaningful to the client, his or her significant others, and the community surrounding the client (e.g., entire classroom, school, and district for children's cases, etc.) (Wolf, 1978). Researchers have emphasized the importance of evaluating social validity in intervention research (e.g., Wolf). However, to date, few researchers have assessed social validity in academic intervention research. Specifically, Wolf stated the following three issues to be addressed when the social validity of an intervention is assessed: (a) the social significance of the intervention goals, (b) the social appropriateness of the intervention procedures, and (c) the social importance of the intervention effects. Schwartz and Baer (1991) classified parties with whom social validity assessment is conducted into the following four groups: (a) direct consumer (e.g., a student); (b) indirect consumer (e.g., a student's significant others such as parents, teachers, etc.); (c) members of the immediate community (e.g., a student's classroom peers, etc.); and (d) members of the extended community (e.g., a student's principal, superintendent in the student's school district).

Researchers have stated different theoretical models of social validity. Eckert and Hintze (2000) reviewed literatures of social validity models and summarized that in general researchers have addressed issues such as complexity, intrusiveness, severity, side effects, and effectiveness of treatments in their assessment of social validity. For example, Witt and Elliott's (1985) model indicates that social validity consists of factors such as acceptability of treatment, treatment use, integrity, and effectiveness. In the model, it is hypothesized that interventions that have high social validity may be selected and used, be implemented with high integrity, and be implemented effectively. The model also indicates that acceptability of treatment is one aspect of a global concept of social validity. Reimers, Wacker, and Koeppl's (1987) model emphasizes that knowledge about the treatment may be associated with treatment acceptability, integrity, and effectiveness. That is, an interventionist may implement an intervention if he or she is knowledgeable about its procedure and prospective outcomes (e.g., an interventionist's knowledge that RP may be used with students with a skill deficit or that it enhanced students' skills), which in turn may enhance integrity and intervention effects. Therefore in the model, researchers may assess how knowledge changes before and after the implementation of intervention, and the interventionist's increase in knowledge may indicate that the intervention is acceptable to the interventionist. Finally, Gresham and Lopez (1996) emphasized that acceptability should be a function of treatment integrity such that high integrity may indicate that the intervention is acceptable to the interventionist; that is, the intervention is implemented with high integrity, maybe because it is acceptable to the interventionist. Additionally, Gresham and Lopez emphasized the use of objective procedures in assessment of acceptability, in addition to subjective procedures (e.g., rating scales), such as semi-structured interview to interventionists regarding intervention effects and social meaningfulness of intervention outcomes and direct observation for integrity assessment.

Traditionally, social validity was assessed as pre-intervention acceptability (e.g., children's acceptability of a group contingency; Elliott, Turco, & Gresham, 1987);

however, social validity has been also assessed as post-intervention acceptability in research (e.g., teachers and students' acceptability of classwide peer tutoring; DuPaul, Ervin, Hook, & McGoey, 1998). In either, social validity results may provide researchers and practitioners with information regarding consumers' likelihood and willingness to participate in or implement the intervention, which in turn aids researchers and practitioners in selecting interventions that may be implemented with high integrity and acceptability. Elliott (1986) also stated that social validity assessment is a methodology to involve children in decision-making (e.g., selection of current or future interventions for themselves or their peers). Inclusion of children (i.e., the direct consumer; Schwartz & Baer, 1991) may promote acceptable decision-making in the RtI model (e.g., the fairness of an intervention and a decision made based on the student's responsiveness to the intervention).

To date, social validity has been assessed in only a few intervention studies in the area of mathematics. For example, DuPaul et al. (1998) implemented classwide peer tutoring in mathematics and reading with elementary school students with attention deficit-hyperactivity disorder and their peers. The study demonstrated that classwide peer tutoring enhanced achievement and was acceptable to the students based on the measure of social validity (e.g., "I would like to have a peer tutor again."). Arra and Bahr (2005) also assessed the social validity of the following three types of mathematics interventions with fourth grade students and teacher-candidates (e.g., university students who majored in elementary education): (a) cognitive (e.g., self-talk and mnemonics), (b) behavioral (e.g., task analysis and tangible reinforcement), and (c) traditional interventions (e.g., drill and practice with little reinforcement). Social validity was assessed before and after the intervention phase. Results indicated that no significant difference in ratings was found before and after the implementation of the interventions or among the three interventions. The authors discussed the results such that the three interventions shared components which were typically used in mathematics classes (e.g., reinforcers such as praise and tangible rewards) and, therefore, no significant difference in acceptability was obtained for the three interventions across time. It should be noted that an intervention package was implemented in the two studies. Social validity may be also assessed for a single intervention component, which will clarify equality or difference in acceptability for potential components.

In summary, in intervention studies, social validity is assessed with regard to the intervention's goal, procedure, and effects (Wolf, 1978). Social validity may be also assessed with diverse consumers of the intervention (e.g., the direct and indirect consumers; Schwartz & Baer, 1991). Given that social validity is associated with intervention effectiveness as well as integrity (Witt & Elliott, 1985), an intervention with high social validity may cause effects on students' academic performance and motivation toward learning (e.g., proactive participation in the intervention). To date, intervention studies in which social validity was assessed have been limited in the area of mathematics (e.g., Arra & Bahr, 2005; DuPaul et al., 1998), and thus, further research that address this issue is warranted.

Self-Efficacy

Self-efficacy is another issue that may be examined in intervention studies. Self-efficacy is an individual's self-perception of capability and competence about current and future performance (Bandura, 1977). Self-efficacy is related to motivation and achievement (e.g., Schunk, 1991). Self-efficacy predicted achievement in general academics in high school students (e.g., grade point average; Caraway, Tucker, Reinke, & Hall, 2003) and mathematics achievement in middle and high school students (Pietsch, Walker, & Chapman, 2003; Stevens, Olivarez, Lan, & Tallent-Runnels, 2004). In mathematics, researchers have found that self-efficacy mediates the effects of ability on mathematics performance (Pajares & Kranzler, 1995; Stevens et al.). Students with SLD demonstrated lower self-efficacy toward their academic performance than those without SLD (Hampton & Mason, 2003; Lackaye, Margalit, Ziv, & Ziman, 2006). Development of interventions that enhance students' self-efficacy in classrooms has been also emphasized (Linnenbrink & Pintrich, 2002). Specifically, researchers and practitioners may implement an intervention with students who are academically at-risk and assess their self-efficacy toward the subject (e.g., mathematics) by administering a brief questionnaire. This may allow researchers and practitioners to predict the students' future academic success (e.g., continuous demonstration of progress) and/or their future engagement in learning the academic subject even after the termination of the intervention.

Although to date, researchers have demonstrated the relationship between self-efficacy and current or future academic performance, they have not well examined how self-efficacy changes as an intervention is being implemented across time. Given that students who exhibit mathematics difficulties may have low motivation or self-efficacy, researchers may examine how self-efficacy changes across time when academically at-risk students receive interventions. Overall, self-efficacy may be a variable that affects students' academic performance and motivation toward learning (cf., Bandura, 1977; Schunk, 1991). However, further research is warranted to examine association between self-efficacy and intervention responsiveness rather than between self-efficacy and a level of academic performance at a single assessment opportunity. Specifically, in intervention studies, researchers may evaluate changes in a student's self-efficacy, along with his or her responsiveness to an intervention and the social validity of the intervention. Results of self-efficacy assessment may provide researchers and practitioners with information regarding likelihood of the student's continuous engagement in learning activities and academic progress even after the termination of the intervention.

Summary for Academic and Mathematics Interventions

In summary, in mathematics, a variety of interventions that address lack of calculation automaticity may be implemented to examine the responsiveness of students who are academically at-risk in the RtI model. Effective interventions should increase students' learning rates and levels of academic skills (Skinner, 1998; Skinner et al., 1996). In mathematics, interventions including practice and feedback components have been demonstrated to enhance students' calculation automaticity (e.g., Bennett & Cavanaugh, 1998, for ICF; Cates, 2005, for RP). Interventions that target different ratios of mastery and instructional levels of facts are also effective in promoting students' fact acquisition (e.g., Cooke et al., 1993, for FI; Skinner, 2002, for interspersal worksheets).

In intervention studies, researchers have also examined changes in related variables that are likely to affect intervention effects, including social validity (Wolf, 1978). Social validity indicates the acceptability of intervention goals, procedures, and effects (Wolf). In the RtI model, along with intervention effects, social validity may be evaluated to determine if intervention process (e.g., procedure) and a decision made based on the intervention effects (e.g., special education eligibility) are meaningful to the diverse consumers of the intervention (e.g., a direct consumer including a student and an indirect consumer including the student's teacher and parent; Schwartz & Baer, 1991). Self-efficacy may be also assessed when interventions are implemented with students who are academically at-risk, given that enhanced self-efficacy may promote these students' future engagement in learning activities and improve their academic progress (cf., Bandura, 1977; Schunk, 1991). The importance of development and implementation of an intervention that enhances students' self-efficacy has been emphasized (Linnenbrink & Pintrich, 2002). However, research is still warranted to examine association between self-efficacy and intervention responsiveness (e.g., whether an intervention enhances both a student's academic performance and self-efficacy).

Finally, in intervention studies, researchers should focus on not only intervention effects, but also selection methodology of the most effective intervention that specifically address a student's academic deficit. One such methodology is a brief experimental analysis (BEA; Daly, Andersen, Gortmaker, & Turner, 2006), in which potential interventions are implemented alternatively and their effects are compared using single subject design methodology. An intervention identified using a BEA may maximize a student's responsiveness to an intervention, which in turn may yield high social validity of the intervention and enhance his or her self-efficacy. Literature regarding BEA and its effects on intervention selection are reviewed in the following section.

Brief Experimental Analysis

BEA is defined as a methodology in which potential interventions are implemented and their effects are compared using single subject methodology (Daly et al., 1997). Researchers and practitioners may conduct a BEA to identify the most effective intervention (or intervention components) for each student, in which they (a) implement the potential intervention components alternatively, (b) subsequently add intervention components until the least sufficient component(s) are identified, or (c) subsequently withdraw intervention components until the least necessary component(s) are identified (Daly, Andersen, Gortmaker, & Turner, 2006). The effects of the selected intervention may be also examined in an extended analysis following the BEA. Effects of BEA have mainly been demonstrated in studies of reading interventions (e.g., Daly & Martens, 1994; Daly, Martens, Hamler, Dool, & Eckert, 1999; Eckert, Ardoin, Daisey, & Scarola, 2000; Eckert, Ardoin, Daly, & Martens, 2002) and less in studies of mathematics and/or writing interventions (e.g., Carson & Eckert, 2003; Duhon et al., 2004). BEA has been also demonstrated to be effective in identification of a student's deficit (i.e., skill or performance) through examination of a student's response to both skill (e.g., direct instruction intended to teach a skill) and performance-based (e.g., contingency reinforcement and other interventions designed to motivate a student) interventions (e.g., Duhon et al., 2004).

Traditionally, in intervention studies, potential interventions have been selected without experimental analyses and/or intervention packages which include the potential components were selected. That is, interventions are determined a priori rather than selected depending upon each a student's specific academic deficit (e.g., fluency versus skill scquisition). Although interventions selected without experimental analysis may be effective in addressing academic deficits in some students, an idiosyncratic intervention identified in BEA may be more effective in addressing the specific deficit of each student. Additionally, in the RtI model, it is crucial to select the best intervention that may maximize students' responsiveness using experimental methodology (e.g., BEA). Researchers have demonstrated that interventions identified in BEA are more effective than those selected without BEA. For example, Carson and Eckert (2003) compared the effects of empirically-identified interventions (i.e., interventions identified in BEA) and student-selected interventions (i.e., interventions that students selected as most effective for them based on their preference and perceptions of the effects of the interventions) on the mathematics calculation performance of elementary-aged students. Specifically, Carson and Eckert implemented performance-based interventions (e.g., contingency reinforcement, goal setting, etc.) using a multi-element design (BC design) in which ATD was embedded in each of the phases. In the first phase, the data of the performance-based interventions and baseline were collected and the most effective intervention was identified (i.e., empirically-derived intervention). In the second phase, only the data of the empirically-derived intervention, the student-selected intervention (i.e., the intervention that the student selected after the first phase), and baseline were collected. Results indicated that the empirically-identified intervention was more effective in enhancing the students' calculation fluency than the student-selected intervention. The results indicate that an empirically-derived intervention may best address a student's academic deficit. Given that many of the students who are referred to the last tier in the RtI model may have a skill rather than a performance deficit, the effects of BEA with these populations should be examined using skill-based interventions.

BEA can be also used to identify a student's type of deficit and an effective intervention that addresses the deficit. Duhon et al. (2004) compared the effects of the following skill and performance-based interventions: (a) direct instruction and (b) contingency reward. An ABC design (A for baseline, B for performance feedback, and C for analysis) was used in this study. The participants were four third through fifth grade students in general education; two students were determined to have a skill deficit (one in writing and one in mathematics) and two students had a performance deficit (both in writing). Baseline data were not collected for one student, because his interventionist mistakenly provided feedback during baseline. The performance feedback phase was introduced to examine whether performance feedback alone was effective on students' performance. In the analysis phase, both direct instruction (e.g., flashcards for mathematics) and contingency reward were implemented with all the students using an ATD. Results indicated that the two students with a skill deficit increased their performance (e.g., the number of digits correct in mathematics) under direct instruction, whereas the two students with a performance deficit increased their performance under contingency reward. Additionally, little data overlap was observed between the two interventions for all the students. This study demonstrated that experimental analysis is effective in the identification of a student's deficit type and a subsequent effective intervention. Additionally, the study demonstrated that matching of deficit and interventions is crucial to enhance the responsiveness of students (e.g., skill-based interventions such as RP for students with a skill deficit). Effects of experimental analysis to examine skill-based (e.g., RP) and performance-based (e.g., feedback and reward) interventions have been also demonstrated in reading (e.g., Eckert et al., 2000).

The reliability of BEA in identification of effective interventions has been demonstrated using a meta-analysis in the area of reading. Burns and Wagner (2008) reviewed 13 intervention studies in reading published from 1994 to 2005, in which BEA was conducted to identify the most effective intervention. The authors reported a high effect size of 2.87 and a high mean percentage of non-overlapping data (PND; Scruggs, Mastropieri, & Casto, 1987) of 81.83% for empirically-derived interventions reviewed in the meta-analysis, which was equivalent to an increase of 30.19 words correctly read per minute (73.0% increase) under an empirically-derived intervention from baseline to BEA. PND indicates the percentage of non-overlapping data points between baseline and intervention phase; high PND indicates the effects of the intervention (i.e., level of desired performance increases or that of undesired performance decreases from baseline to the intervention phase, yielding high PND). The results also indicated that some of the skill-based interventions had high PND (e.g., 100% for repeated reading), whereas performance-based interventions had low PND (e.g., 50% for incentive). These facts may indicate that skill-based interventions are more effective in enhancing the skills of students with reading deficits than performance-based interventions. This may be also true for students with mathematics deficits, since lack of automaticity is a common aspect of these students and thus interventions should address skill acquisition (e.g., automaticity development) rather than motivation enhancement (Hasselbring et al., 1988). Overall, BEA may be a reliable method to identify an idiosyncratic intervention in reading. Despite lack of comprehensive meta-analytic reviews regarding BEA in mathematics, the reliability of BEA may be also assumed in mathematics.

In summary, an intervention identified based on BEA may maximize the likelihood of a student's responsiveness in the RtI model. However, the dearth of studies that have examined this issue in mathematics is problematic. For academically at-risk students, potential skill-based interventions (e.g., RP, etc.) may be implemented in BEA and their effects may be compared. Additionally, the goal of the intervention for an academically at-risk student may be to enhance the student's skills on his or her instructional level materials, maintain the skills on the instructional materials, and generalize the skills to his or her grade level materials. Therefore, generalization and maintenance of intervention effects may be also examined when implementing academic interventions in the RtI model. Researches on generalization and maintenance are reviewed in the following section.

Generalization and Maintenance

Generalization refers to a phenomenon in which a behavior addressed through an intervention is likely to occur under different stimuli (e.g., people, places, and time) or causes other similar behaviors (Baer, Wolf, & Risley, 1968). Generalization should be programmed into an intervention to increase the likelihood that it will deliberately occur (Stokes & Baer, 1977). Additionally, maintenance is defined as a phenomenon in which a behavior addressed through an intervention is likely to be demonstrated during and/or after termination of the intervention. Generalization and maintenance are also components of the Instructional Hierarchy (i.e., acquisition, fluency, generalization, and adaptation) (Daly, Lentz, & Boyer, 1996; Haring & Eaton, 1978). In Instructional Hierarchy, students who have acquired fluency in skills are expected to generalize and maintain the skills under the same or similar instructional stimuli. For academically at-risk students, academic interventions that address stages of Instructional Hierarchy (interventions that address not only acquisition and/or fluency, but also generalization) may be implemented to enhance their skills on both instructional and grade level materials. In mathematics, students' skills may generalize across different grade levels, skills, and settings. In intervention studies, students may generalize skills, which they have acquired through interventions, on academic stimuli that they have not learned (e.g., academic skills that were not targeted in the interventions). Specifically, students may apply skills from his or her instructional level materials (e.g., second grade materials including 2-digit plus 2-digit addition) to his or her grade level materials (e.g., third grade materials including 3-digit plus 3-digit addition) (i.e., generalization across grade levels).

On the other hand, students are expected to maintain skills that they acquired in the intervention during and/or after it is terminated. Thus, in intervention studies, researchers may examine generalization and maintenance of skills that are targeted in an intervention while examining the intervention effects. Effective interventions should also assist students in acquiring skills targeted in multiple stages in Instructional Hierarchy (e.g., instructional skills targeted in the acquisition and fluency stages, generalization skills in the generalization stage).

Generalization and maintenance have been examined predominantly in reading intervention literature. For example, Bonfiglio, Daly, Martens, Lin, and Corsaut (2004) examined the effects of the following three types of reading interventions and their generalization and maintenance with one third grade student: (a) performance-based treatment (e.g., contingency reward), (b) skill-based treatment (e.g., repeated reading), and (c) combined performance-based and skill-based treatment (containing all components of both interventions). Six passages consisted of three easier passages (Passages 1-3) and three more difficult passages (Passages 4-6). Two ABABA designs were used in parallel across two passage types (i.e., easier and more difficult passages). For each session, all six passages (Passages 1-6) were administered in a counterbalanced order such that one of the three easier passages and one of the more difficult passages were examined under one of the three interventions (e.g., Passages 1 and 4 under the skill-based treatment, Passages 2 and 5 under the performance-based treatment, and Passages 3 and 6 under the combined treatment). In the study, all six passages were administered during each baseline. During a first intervention phase, the three interventions were implemented on the three easier passages using the simultaneous-treatment design (Hayes et al., 1999), in which the three interventions were implemented in each session with a different order; no interventions were implemented on the three difficult passages. On the other hand, during a second intervention phase, the three interventions were implemented on the three difficult passages using the simultaneous treatment design; no interventions were implemented on the three easier passages this time. Results showed that the student's oral reading fluency improved on easier passages from baseline to the first intervention phase targeting easier passages. During a second baseline, the student's oral reading fluency slightly declined on easier passages, but slightly improved on more difficult passages administered as generalization passages (i.e., generalization from easier to more difficult passages). During the second intervention phase targeting difficult passages, the student's oral reading fluency improved on difficult passages. During the last baseline, the student's oral reading fluency slightly declined on more difficult passages, but it significantly improved on easier passages despite the lack of interventions on these passages in the previous phase (i.e., maintenance on easier passages). This study is one example of generalization from easier to more difficult passages and maintenance in oral reading fluency. However, few studies have been examined generalization in mathematics in a similar rigorous manner.

In summary, intervention research in mathematics should address generalization and maintenance. Given low levels of students with SLD in mathematics, researchers should examine generalization from instructional level materials (e.g., instructional worksheets from lower level materials) to their grade level materials (e.g., worksheets at the frustrational level and at students' grade level) and maintenance on both types of materials across time.

Statement of the Problem

As reviewed, the RtI model has been presented as an alternative model for SLD identification. However, further research is needed to examine the validity of the RtI model in identification of SLD in mathematics calculation. Therefore, the current study addressed the following five issues. First, responsiveness to intervention should be examined using a single subject design, given that responsiveness is idiosyncratic. Second, an intervention should be experimentally identified depending on each individual's deficit type (e.g., skill deficit), using an experimental analysis. Third, responsiveness to intervention should be examined for both academically at-risk students in general education and students with SLD in mathematics calculation in special education. This may allow examination of the ability of RtI methodologies to identify responders and nonresponders in both classification categories. Fourth, generalization of skills from students' instructional levels to their grade levels and maintenance of the skills for both levels should be examined. Although students diagnosed with SLD in mathematics calculation may respond to lower grade level materials, interventions may be expected to allow generalization of their skills to grade level materials. Finally, students' social validity of mathematics intervention and self-efficacy should be assessed within the RtI model, which may assist in the measurement of the acceptability of the intervention selected for the student.

Purpose of the Study

The purpose of this study was to compare the responsiveness to intervention of elementary school students with and without SLD in mathematics. Given that responsiveness is idiosyncratic, each student's responsiveness was analyzed using a single subject design rather than a group design. Data were examined visually for changes in level, trend, and variability across the data points. For each student, the most effective intervention was identified using a brief experimental analysis. Subsequently, the intervention was implemented for approximately 6 weeks to examine his or her intervention responsiveness (e.g., weekly slopes of progress). Mean intervention responsiveness was also compared between students with and without SLD. Additionally, generalization from intervention materials to the student's grade level materials and maintenance on both types of materials were examined. Follow-up data were collected on instructional and generalization materials one week after termination of the intervention. Furthermore, social validity and self-efficacy were assessed with each student. Mean social validity and self-efficacy ratings were also compared between students with and without SLD, respectively.

Hypotheses

To accomplish the purpose of this study, the following two hypotheses were examined.

1. Students, with and without SLD in mathematics calculation, would demonstrate responsiveness when an intervention was identified through a brief experimental analysis and implemented for approximately 6 weeks.

2. Students, with and without SLD in mathematics calculation, would demonstrate satisfaction and self-efficacy following implementation of the intervention that best addressed their skill deficit.

CHAPTER II

METHODOLOGY

In this chapter, the methods and procedures that were used to implement mathematics interventions, collect data, and analyze the data are stated. Specifically, the following issues are presented: (a) participants and selection criterion, (b) materials, (c) independent variables, (d) dependent variables, (e) procedures, and (f) design and analysis.

Participants and Selection Criteria

Participants were selected from general and special education classes in five elementary schools in a large community in the southern United States. The district where data were collected included approximately 50 elementary schools. Based on the most recent demographic and educational data in the district, 61.9% of the pre-kindergarten to twelfth grade students received federal free or reduced price lunch program (61.2% for the state); 46.8% and 47.3% of the third and fourth grade students, respectively, scored within the lowest achievement level on statewide achievement tests in mathematics (37.3% and 37.9% for the state, respectively); and 18.3% of the students aged from 3 to 21 in special education met the state criteria for Specific Learning Disability (SLD) (33.0% for the state) (Louisiana Department of Education, 2006, 2007).

The Response to Intervention (RtI) model was introduced in the district the previous year before this study was conducted. Prior to the introduction of the RtI model, prereferral interventions were implemented with students who had academic and/or behavioral problems, before they were referred for comprehensive assessment (e.g., standardized intelligence and/or achievement tests); however, the tiers of RtI were not applied in these interventions.

Participants Selection

In this section, the following sequence of participant selection is outlined: (a) district curriculum-based measurement (CBM) screening (b) initial selection of general education students, (c) initial selection of special education students, (d) pre-treatment assessment, (e) skill deficit assessment, (f) consent and assent procedure, (g) teacher interview, and (h) selection of the final participants.

District CBM screening. During the year of this study and the previous year, academic screening was conducted with all students in general education and some students in special education (e.g., students with SLD, Other Health Impairment, etc.) in elementary schools three times a year (fall, winter, and spring) using CBM in the area of reading and mathematics. In the CBM screening, mixed-skill worksheets were administered for different time durations (e.g., 2 minutes for second grade; 3 minutes for third grade). Additionally, during the year of this study, the consultation team at the local university developed the materials (e.g., mathematics worksheets) used in the screening based on Louisiana Grade Level Expectations (GLE; Louisiana Department of Education, 2004). The consultation team also assisted teachers in conducting CBM screening and identifying students who were academically at-risk, suggested individual and classwide interventions, and monitored intervention progress. In the state where the district existed, SLD was identified using an achievement-achievement discrepancy model, in which a discrepancy of 15 points (i.e., one standard deviation) or more between the student's strength and weakness areas on a standardized achievement test yielded an identification of SLD in the area(s) of concern.

Initial selection of general education students. In general education classes, third grade students who were identified based on teacher nomination as academically at-risk in mathematics calculation and who attended school during the fall semester were selected for participation in this study. Teachers selected students in their class whose level was within frustrational level based on fall CBM screening data and whose academic functioning was at least one grade below third grade expectations. Statewide achievement test scores were unattainable for third grade students. Therefore, for the third grade participants in general education, CBM scores for the third grade collected in the district fall screening were used as one of the selection criteria. Additionally, all general education students had to exhibit a skill deficit (i.e., lack of skills) in mathematics calculation rather than a performance deficit (i.e., lack of motivation) to be included in the study. Skill versus performance deficit was determined by a skill deficit assessment (see procedures section for details).

Initial selection of special education students. Third and fourth grade students identified with SLD in mathematics calculation were selected from special education classes based on teacher nomination. Special education teachers selected a third grade student whose CBM scores in the fall screening were within frustrational level or a fourth grade student whose statewide achievement test (i.e., Louisiana Educational Assessment Program test; Louisiana Department of Education, 2005-2006) scores were within the lowest level (i.e., unsatisfactory level, indicating below 56% achievement levels) for the previous year and for whom the teachers had the most concern regarding mathematics calculation skills. Thus, these students had a skill deficit in mathematics calculation.

Pre-treatment assessment. For the selected general and special education students, curriculum-based assessment (CBA) was used to confirm their instructional level. During CBA, for each student, three multiple-skill worksheets at each grade were administered. The number of digits correct per minute (DCPM) and errors per minute (EPM) were calculated for each worksheet. Median DCPM and EPM were also calculated for the grade (see procedures section for details regarding determination of a student's instructional level). Students whose instructional levels were below first grade (e.g., students who have difficulty in number sense or counting) were excluded from the current study. To prevent confounding results, students who were identified with severe developmental disorders (e.g., Autism, Asperger's Syndrome, Mental Retardation) or severe emotional or behavioral disorders (e.g., Depression, Attention Deficit/Hyperactivity Disorder) were excluded. Students who were currently receiving the mathematics interventions to be used in the study were also excluded.

Skill deficit assessment. For a student whose instructional level was identified as first grade or above based on the CBA, a skill deficit assessment was also conducted to determine whether the student had a skill or performance deficit at the instructional level. Students who had a performance deficit based on results of the skill deficit assessment (e.g., exhibited motivational difficulties rather than skill-base difficulties) were excluded from this study. For students who had a skill deficit, known and unknown facts at the instructional level were also assessed using approximately 100 flashcards at that level; facts responded to correctly within 3 seconds were considered as known facts. The percentage of known facts was also calculated as an indicator of the student's initial level of performance, along with the level identified during the CBA (i.e., frustrational, instructional, or mastery). This allowed examination of association between initial level of performance and intervention responsiveness for all students as well as additional monitoring of progress across the study for students who responded to the FI intervention.

Consent and assent procedure. The researcher obtained the permission of the superintendent in the district. Once Institutional Review Board (IRB) approval was obtained (see Appendix C), the researcher met with each teacher, gave the teacher the letter indicating the purpose of the study, and asked the teacher to identify potential students for inclusion in the study and to send a parent consent form to the parents of the students. A parent consent form was sent to the parents of 8 students in general education and 10 students in special education at eight elementary schools in the district; the parents of 3 students in general education did not return a consent form to school. Thus, parental consent was obtained for 5 students in general education and 10 students in special education. For students for whom parental consent was obtained, child assent was also obtained. Child assent was obtained from all of these students except for one student with SLD; the student indicated on his assent form that he would not like to participate in this study and thus was excluded.

Teacher interview. Once consent and assent were obtained, each student's teacher was interviewed regarding his or her concerns about the student's achievement (e.g., instructional level) and classroom environment (e.g., instructional method) (see Appendix D). Teacher interview was conducted at the end of the fall semester for some students and at the beginning of the spring semester for the other students. The teacher interview revealed that one student with SLD had a medical diagnosis of Attention Deficit/Hyperactivity Disorder based on his evaluation report in the district; thus, he was excluded from the study.

Selection of the final participants. During the subsequent phases, CBA revealed that one student without SLD had a performance deficit and three students with SLD had skills below the first grade; thus, these four students were excluded. During the experimental analysis or intervention phase, one student without SLD and two students with SLD were excluded from this study due to their frequent absence from school or transition to another school.

As such, three students without SLD and three students with SLD were included in this study. The information obtained in teacher interview for the final participants in this study is summarized below. (See Table 2.1 for the final 6 students' demographic information. See Tables 2.2 and 2.3 for information for SLD identification and special education placement for the students with SLD.)

Demographics of the Participants in General Education

Jonathan. Jonathan was an 8-year-old Caucasian male student in a third grade general education class. He had not repeated any previous grades. Jonathan's fall CBM screening data were 0.7 DCPM for the third grade and 5 DCPM for the second grade (see Table 2.1 for summary data for all general and special education students). The areas targeted in his mathematics class at the time of the interview included multiplication facts. No mathematics intervention was implemented with Jonathan prior to or at the time of the teacher interview. CBA revealed that his instructional level was at the first grade one-digit plus one-digit addition sums to 10 (*Mdn* = 10 DCPM, *Mdn* = 1 EPM) (see Table 2.1). He knew 82.8% of the facts for the instructional level at the time of the CBA.

Jamario. Jamario was an 8-year-old African American male student in a third grade general education class. He was in the same classroom as Jacklyn. Jamario had not repeated any previous grades. Jamario's fall CBM screening data were 3.7 DCPM for the third grade and 5.5 DCPM for the second grade (see Table 2.1). No mathematics intervention was implemented with Jamario prior to or at the time of the teacher interview. CBA revealed that his instructional level was at the second grade for mixed-skill worksheets (*Mdn* = 16 DCPM, *Mdn* = 0 EPM) and his instructional skill was one-digit plus one-digit addition sums to 18 (*Mdn* = 15 DCPM, *Mdn* = 0 EPM) (see Table 2.1). He knew 69% of the facts for the instructional level at the time of the CBA.

Table 2.1 Demographic Information of the Participants

Student	Age	Grade	District CBM/ statewide achievement test scores	Repeated grades	Intervention	Instructional skill
			General Education			
Jonathan	8	3	5.0 DCPM 0.7 EPM	N/A	RP	1-digit plus 1-digit sums to 10
Jamario	8	3	5.5 DCPM 3.7 EPM	N/A	FI	1-digit plus 1-digit sums to 18
Jacklyn	8	3	4.5 DCPM 0.7 EPM	N/A	ICF	1-digit plus 1-digit sums to 10
			Special Education			
Kathy	12	4	47% (4th) 33% (3rd)	1st-3rd 4th	RP	1-digit plus 1-digit sums to 10
Kaley	12	4	36% (4th) 28% (3rd)	1st, 4th	FI	1-digit plus 1-digit sums to 10
Tyson	12	4	55% (4th) 16% (3rd)	1st, 4th	ICF	1-digit plus 1-digit sums to 18

Note. Kathy may have repeated one of the grades from first through third. CBM = curriculum-based measurement; DCPM = digits correct per minute; EPM = errors per minute; RP = repeated practice, FI = folding-in technique; ICF = immediate corrective feedback.

Table 2.2 Information for Specific Learning Disability (SLD) Identification and Special Education Placement

Student	Age at SLD diagnosis	Grade at SLD diagnosis	Years with SLD diagnosis	SLD categories	Special education placement
Kathy	7	1st	5	MC, BR, RC, WE	Mathematics, Reading, English Language Arts
Kaley	10	3rd	2	MC, MR	Mathematics
Tyson	11	4th	1	MC, BR, RC, WE	Mathematics, Reading, English Language Arts

Note. SLD = Specific Learning Disability; MC = Mathematics Calculation; MR = Mathematics Reasoning; BR = Basic Reading; RC = Reading Comprehension; WE = Written Expression.

Table 2.3 Standardized Test Results for Students Diagnosed with SLD in Mathematics Calculation

Student	WJ-III ACH score on Math Calc	WJ-III ACH grade-equivalent level
Kathy	*	<K.2
Kaley	67	1.7
Tyson	66	1.0

Note. WJ-III ACH = *Woodcock-Johnson III, Test of Achievement* (Woodcock, McGrew, & Mather, 2001).
For Kathy, * indicates that a raw score on Math Calc was 0; thus, a standardized score was not calculated.

Jacklyn. Jacklyn was an 8-year-old African American female student in a third grade general education class. She had not repeated any previous grades. Jacklyn's fall CBM screening data were 0.7 DCPM for the third grade and 4.5 DCPM for the second grade (see Table 2.1). The areas targeted in her mathematics class at the time of the interview included multiplication facts and division facts. No mathematics intervention was implemented with Jacklyn prior to or at the time of the teacher interview. CBA revealed that her instructional level was at the first grade one-digit plus one-digit addition sums to 10 (*Mdn* = 13 DCPM, *Mdn* = 0 EPM) (see Table 2.1). Jacklyn knew 79.7% of the facts for the instructional level at the time of the CBA.

Demographics of the Participants in Special Education

Kathy. Kathy was a 12-year-old Caucasian female student in a fourth grade resource class. It should be noted that in schools, during resource classes, students receive instructions in special education classes only in their weakness subject areas, whereas during self-contained classes, students receive instructions in special education in all subject areas. She had a ruling of SLD in Mathematics Calculation (MC), Basic Reading (BR), Reading Comprehension (RC), and Written Expression (WE) with a strength in Mathematics Reasoning (MR) (the discrepancy between her academic strength and weakness areas yielded her SLD identification in the weakness area) (see Tables 2.1 to 2.3). She was initially identified with Speech/Language Impairments at the age of 4 years old (kindergarten) and then as SLD at the age of 7 years old (first grade). The referral for the SLD identification was based on her difficulty in word identification. A peer tutoring intervention using flashcards that targeted word identification was implemented with her as a prereferral intervention. In the SLD evaluation, her standard score on the MC subscale in the *Woodcock-Johnson III, Test of Achievement* (WJ-III ACH; Woodcock, McGrew, & Mather, 2001) was not available due to her raw scores of 0 on the Math Calculation and Fluency subtests under the MC subscale (K.2 grade level) (see Tables 2.3). Kathy repeated the fourth grade due to her failure in mathematics and English Language Arts. Although her educational records indicated only repetition of the fourth grade, her evaluation and triennial re-evaluation reports in the district indicated that she spent four years from the first through third grades. This indicates that she might have repeated one of the grades from first through third or might have been out of school for a certain period. She did not take the fall CBM screening test due to her absence on the screening day; however, her level on a statewide achievement test in mathematics was approximately 47% achievement level (unsatisfactory) for the initial fourth grade and approximately 33% achievement level (unsatisfactory) for the third grade. Kathy was receiving speech therapy in school; however, her teacher stated that her speech difficulty did not affect her classroom performance in mathematics. She attended mathematics, reading, and English Language Arts classes in special education (approximately 10 students) and science and social studies in general education. The skills targeted in her mathematics class at the time of the interview included multiplication facts and division (e.g., one-digit divided by one-digit with and without a remainder). The teacher used accommodations including individual instructions as needed. Kathy was not attending before or after-school programs. No mathematics intervention was implemented with her prior to or at the time of or prior to the teacher interview. CBA revealed that Kathy's instructional level was at the first grade one-digit plus one-digit addition sums to 10 (*Mdn* = 19 DCPM, *Mdn* = 3 EPM) (see Table 2.1). She knew 82.8% of the facts for the instructional level at the time of the CBA.

Kaley. Kaley was a 12-year-old African American female student in a fourth grade inclusion class. She had a ruling of SLD in MC and MR with strengths in RC and WE (see Tables 2.1 to 2.3). She was initially identified with SLD at the age of 10 years old (third grade). The referral for the SLD evaluation was based on her difficulty in mathematics calculation (e.g., addition and multiplication). A daily 10-minute mathematics intervention that targeted addition or multiplication (e.g., individual instruction) was implemented with Kaley as a prereferral intervention. In the SLD evaluation, her standard score on the MC subscale in the WJ-III ACH was 67 (1st percentile; 1.7 grade level) (see Table 2.3). She repeated the first grade due to her failure in mathematics, reading, and English Language Arts and fourth grade due to her failure in mathematics and English Language Arts. She also failed in mathematics, reading, and English Language Arts in the second grade, but she proceeded to the third grade. A fall CBM screening was not administered to the students in special education at her school; however, her level on a statewide achievement test in mathematics was approximately 36% achievement level (unsatisfactory) for the initial fourth grade and approximately 28% achievement level (unsatisfactory) for the third grade. Kaley did not receive any other services in school (e.g., speech therapy). She attended all classes in general education except for mathematics; she attended a mathematics class in special education (approximately 10 students). The skills targeted in her mathematics class at the time of the interview included multiplication facts and division (e.g., one-digit divided by one-digit with and without a remainder). The teacher used accommodations including individual or small group instructions. Kaley was not attending before or after-school programs. No mathematics intervention was implemented with her prior to or at the time of the teacher interview, except for the prereferral intervention. CBA revealed that Kaley's instructional level was at the first grade one-digit plus one-digit addition sums to 10 (*Mdn* = 16 DCPM, *Mdn* = 1 EPM) (see Table 2.1). She knew 76.2% of the facts for the instructional level at the time of the CBA.

Tyson. Tyson was a 12-year-old African American male student in a fourth grade inclusion class. He had a ruling of SLD in MC, BR, RC, and WE with a strength in MR (see Tables 2.1 to 2.3). He was initially identified with Speech/Language Impairments at the age of 8 years old (first grade) and then as SLD at the age of 11 years old (the first time of the fourth grade). The referral for the SLD evaluation was based on his difficulty in reading (lack of reading accuracy and fluency). A word acquisition intervention was implemented with him for 10 days as a prereferral intervention. In the SLD evaluation, his standard score on the MC subscale in the WJ-III ACH was 66 (1.0 grade level) (see table 2.3). He repeated the first grade due to his academic failure in the areas of reading, mathematics, and English Language Arts and the fourth grade due to his academic failure in the areas of mathematics and English Language Arts. He did not take the fall CBM screening test due to transition from one school to another in the same district; however, his level on a statewide achievement test in mathematics was approximately 55% achievement level (unsatisfactory) for the initial fourth grade and approximately 16% achievement level (unsatisfactory) for the third grade. Tyson was receiving speech therapy in school; however, his teacher stated that his speech difficulty did not affect his classroom performance in mathematics. During the fall semester, he attended all classes in general education with the assistance of his special education teacher except for English Language Arts class (special education). However, during the spring semester, he attended mathematics, reading, and English Language Arts in special education (approximately 10 students) and science and social studies in general education due to parental request. The skills taught in his mathematics class at the time of the interview included multiplication facts and division (e.g., two-digit divided by one-digit with and without a remainder). His special education teacher used accommodations such as an approximately 15-minute small group instruction consisting of 2 to 3 students and individual instructions as needed in both general and special education classes during the fall and spring semesters. Tyson was not attending before or after-school programs. No mathematics intervention was implemented with him prior to or at the time of the teacher interview. CBA revealed that Tyson's instructional level was second grade (*Mdn* = 26 DCPM, *Mdn* = 3 EPM) for mixed-skill worksheets and his instructional skill was one-digit plus one-digit addition sums to 18 (*Mdn* = 32 DCPM, *Mdn* = 0 EPM) (see Table 2.1). Tyson knew 80% of the facts for the instructional level at the time of the CBA.

Materials

The following sections describe the instruments used in this study: (a) mathematics calculation worksheets, (b) supplies used in administering the worksheets, and (c) the *Student Social Validity Checklist*.

Mathematics Calculation Worksheets

Five types of worksheets were used in this study: (a) CBA, (b) instructional, (c) intervention monitoring, (d) maintenance (i.e., a cold worksheet), and (e) generalization. All of these worksheets were generated from the *Mathematics Worksheet Factory* (Schoolhouse Technologies, Inc., 2006) and were based on Louisiana Mathematics Benchmarks for each grade (GLE; Louisiana Department of Education, 2004). The researcher retyped all worksheets to create worksheets in the same format (e.g., font size). For all students, all of these worksheets included 30 calculation problems, which were considered a typical number of problems that were administered in third and fourth grade classes (See Appendix E). Given that students who are academically at-risk in general education and students with SLD in special education are likely to have deficits in addition (e.g., basic mathematics facts; Hasselbring, Goin, & Bransford, 1988), only addition problems were included in the worksheets used in this study (e.g., 30 problems of only one-digit plus one-digit addition sums to 10). The purpose of each type of worksheet is described below.

CBA worksheets. For each student, CBA worksheets were used to determine each student's instructional level based on the CBM criterion (Deno & Mirkin, 1977, cited in Shapiro & Lentz, 1986; see Appendix B). Each CBA worksheet included mixed instructional level problems at the grade. For students whose instructional level was identified as second grade, CBA worksheets at each benchmark at second grade (i.e., single skill such as one-digit plus one-digit sums to 18) were also administered to determine the student's instructional benchmark.

Instructional worksheets. Instructional worksheets were used to teach calculation skills during the following two interventions: (a) repeated practice (RP) and (b) immediate corrective feedback (ICF) (materials used in these interventions are stated in this section; however, procedures for each intervention are stated in the procedures section). In folding-in technique (FI), flashcards rather than instructional worksheets were used to teach calculation skills (procedures for flashcards are also stated in this section). Instructional worksheets included single-skill problems at the student's instructional level.

Intervention monitoring worksheets. *Intervention m*onitoring worksheets were used to monitor progress after practicing calculation problems during intervention. An intervention monitoring worksheet was developed by re-ordering all problems on the corresponding instructional worksheet used for teaching skills during intervention. For FI, the researcher also created an intervention monitoring worksheet with 70% known facts and 30% unknown facts by placing these facts in a randomized order on each worksheet; these modified intervention monitoring worksheets were used in folding-technique rather than typical intervention monitoring worksheets (i.e., worksheets with a random ratio of known and unknown facts used in the other two interventions).

Maintenance worksheets. Maintenance worksheets (i.e., cold worksheets) were used to assess maintenance of the skill targeted in the intervention. Maintenance worksheets included single-skill problems at the student's instructional level.

Generalization worksheets. Generalization worksheets were used to assess generalization of the skill targeted in the intervention to grade level skills. Generalization worksheets targeted multiple skills at the student's grade including the skill of focus in the corresponding instructional worksheet. For example, if an instructional worksheet that included one-digit plus one-digit addition sums to 10 (i.e., first grade skill) was administered to a third-grade student, a generalization worksheet that included three-digit plus three-digit addition with or without regrouping (i.e., third grade skills) was also administered to the student after he or she has completed the instructional worksheet.

Flashcards

Flashcards of mathematics facts (3 inch x 5 inch) were also used in the implementation of FI. A mathematics calculation problem was printed on the front of each card and the answer was printed on the back of the card. Given that students were expected to acquire automaticity in solving these basic mathematics calculation problems, mathematics facts printed on cards included instructional level problems of addition for each student.

Other Materials

Stopwatches and pencils were used to implement the curriculum-based procedures (see procedures section).

Student Satisfaction and Self-Efficacy Scales

The *Student Social Validity Checklist* (see Appendices H.1 to H.4) was administered to assess a student's satisfaction with intervention and to determine if their sense of self-efficacy improved following intervention. Intended to measure the student's satisfaction with the provided intervention and his or her perception of self-efficacy in solving presented mathematics problems, the researcher developed the *Student Social Validity Checklist* based on the *Children's Intervention Rating Profile* (CIRP; Witt & Elliott, 1985) and the *Self-Perception Profile for Children* (Harter, 1985). This measure was administered following each condition within the study. The two measures, the CIRP and the *Self-Perception Profile for Children, Scholastic Competence* scale, which were used to develop the student social validity checklist, are briefly described below followed by a brief section describing the development of the checklist and validity information about the new measure.

CIRP

Witt and Elliott (1983) developed the CIRP, a social validity checklist used to assess children's social validity of behavioral interventions (Elliott, Turco, & Gresham, 1987; Turco & Elliott, 1986). The CIRP is a truncated version of the *Intervention Rating Profile-20* (IRP-20; Witt & Martens, 1983), a 20-item, five-factor scale that assesses the effectiveness of a behavioral intervention and its acceptability to teachers. The CIRP included 7 items each of which consists of a 6-Likert-point scale ranging from *I agree* to *I do not agree*. The CIRP has adequate reliability (e.g., high internal consistency) and validity (e.g., construct validity based on its one-factor model) (Elliott, 1986). The CIRP is a one-factor scale that measures the acceptability of intervention procedures (Elliott), which is one component of Wolf's (1978) social validity model (i.e., acceptability of goals, procedures, and effects). Specifically, the CIRP includes items regarding the following aspects of the intervention: (a) fairness (e.g., "This method used to deal with the behavior problem was fair"); (b) acceptability (e.g., "I like the method used for this child's behavior problem"); and (c) effectiveness (e.g., "I think that the method used for this problem would help this child do better in school"). Each item is rated on a 6-point Likert-type scale ranging from 1= "*I agree*" to 6= "*I do not agree*." A total score (range = 7-42) yields the student's acceptability for the intervention. A modified version of the CIRP has been used to assess the social validity of academic interventions in previous studies (e.g., Arra & Bahr, 2005). Arra and Bahr reported that the modified CIRP used for assessing children's perceptions of different types of mathematics interventions in their study (e.g., a cognitive instructional training intervention) showed acceptable internal consistency (Cronbach's alpha range = .55 - .79). Given that intervention procedures rather than outcomes may affect a student's current performance, as well as future use of the intervention, students' perception of intervention procedures was assessed in this study.

Self-Perception Profile for Children

Harter's *Self-Perception Profile for Children* (Harter, 1985) is a self-administered inventory designed to assess children's perceived competence in the six domains (e.g., Scholastic Competence, Social Acceptance, etc.) and includes 36 items. Each item consists of two sentences with opposite meanings that are connected by the conjunction "BUT" (e.g., "Some kids feel that they are very good at their schoolwork BUT other kids worry about whether they can do the schoolwork assigned to them."). Response to items is a two step process: (a) first, the examinee makes a selection (e.g., "*Which kind of kids are most like you?*") between two opposing statements, and then (b) the examinee determines the degree of validity of their selected statement in a response to a follow-up question (e.g., "*This is Really True for me*" or "*This is Sort of True for me*"). Responses are assigned a score on a 4-point Likert-type scale (e.g., 4= "*Really True for me*" and

3= "*Sort of True for me*" for positive sentences such as "Some kids feel that they are very good at their schoolwork", whereas 1= "*Really True for me*" and 2= "*Sort of True for me*" for negative sentences such as "Some kids are pretty slow in finishing their schoolwork." High scores indicate high self-perception. A total score (range = 4-24 for the Scholastic Competence subscale; range = 4-144 for the entire profile) on each scale (e.g., Scholastic Competence) yields the student's self-perception (e.g., self-efficacy) in that area. Scores can be combined to yield a total score for self-perception. The *Self-Perception Profile for Children* has adequate reliability (internal consistency) and validity (construct validity based on factor analysis; Harter, 1985; Muris, Meesters, & Fijen, 2003).

Overall, the research-based scales reviewed in this section can be modified to develop scales that assess students' satisfaction with intervention procedures and changes in their sense of self-efficacy when implementing interventions. Development of the satisfaction and self-efficacy scales used in the present study is outlined below.

Development of the Student Social Validity Checklist

The 10-item *Student Social Validity Checklist* used in this study consists of the following two scales: (a) Satisfaction Scale (5 items; see Appendix F to H), and (b) Self-Efficacy Scale (5 items; see Appendix I).

Satisfaction scale. To develop the Satisfaction Scale, five of the seven items from the CIRP (items 1, 2, 5, 6, 7) were modified to accommodate the nature of the student-tasks in this study (e.g., "The method used to deal with the behavior problem was fair" was changed to "This intervention to improve my math skills was fair"); two items (items

“3: The method used to deal with the behavior may cause problems with the child’s friends” and “ 4: There are better ways to handle this child’s problem than the one described here”) were excluded, given that each of these items measures similar perception targeted by one of the five selected items. Given that the CIRP is a one-factor scale which measures a global concept of intervention acceptability, some items may measure similar perceptions. For example, items 3 and 5 (“The method used by this teacher would be a good one to use with other children”) measure appropriateness of the use of the intervention with the child’s peers; therefore, a positive item (i.e., item 5) was selected for this study. It should be noted that in previous research, adding (e.g., *Behavior Intervention Rating Scale*, VonBrock & Elliott, 1987) or subtracting (e.g., *Abbreviated Acceptability Rating Profile*; Tarnowski & Simonian, 1992) some items from the IRP-15, the original acceptability scale of the CIRP, did not affect the psychometric properties of the modified IRP-15 (e.g., internal consistency of .98 for the *Abbreviated Acceptability Rating Profile*; Tarnowski & Simonian). Additionally, modification from behavioral to academic terms on items on the CIRP did not affect the psychometric properties of the modified CIRP in previous intervention studies (e.g., acceptable internal consistency; Arra & Bahr, 2005). Based on these previous modification findings, in this study, it was assumed that modification and truncation (exclusion of two items) of the CIRP would not affect the psychometric properties of the Satisfaction Scale.

On the Satisfaction Scale, items consisted of a statement (e.g., “This intervention to improve my mathematics skills was fair”) followed by a 3-point Likert-type rating

(3= "*True*," 2 = "*Somewhat true*," and 1= "*Not true*"). The rating scale used face pictures (happy, neutral, and unhappy faces) rather than numbers to represent the degree of satisfaction (e.g., a happy face indicating 3), which was considered appropriate for the age of students in the study. These items comprised the Satisfaction Scale and provided a Total Satisfaction Score ranging from 5 to 15. A mean score was calculated for each student for this scale, given that a mean score indicates the student's average perception across the items.

Self-efficacy scale. To develop the Self-Efficacy Scale, five of the six items from the *Self-Perception Profile for Children, Scholastic Competence* scale (items 1, 7, 13, 19, 31) were selected and also modified to accommodate this study by selecting only a positive statement for each item (e.g., "Some kids feel like they are just as smart as other kids their age BUT other kids aren't so sure and wonder if they are as smart" was changed to "I feel I am as smart as my classmates"); item 25 was excluded, given that it measures students' perception of competency on a task in a specific setting ("Some kids do very well at their classwork") rather than general academic tasks that may be common in diverse settings (e.g., "Other kids can always figure out the answers"). However, modified versions of Harter's profile or scales in the profile have not been developed in previous studies. It should be noted that each of the six items in the *Scholastic Competence* scale on Harter's profile showed almost equal median loading in the factor analysis reported in the manual (range = .59-.65; Harter, 1985), which may indicate that the six items measures the same concept at almost an equal degree. Thus, it was assumed that exclusion of the item 25 (loading *Mdn* = .65) did not affect the psychometric properties of the Self-Efficacy Scale.

On the Self-Efficacy Scale, items consisted of a statement (e.g., "This intervention to improve my mathematics skills") followed by a 3-likert-type rating (3= "*True*," 2 = "*Somewhat true*," and 1= "*Not true*"). The rating scale used face pictures (happy, neutral, and unhappy faces) with each face representing the degree of self-efficacy (e.g., a happy face indicating 3), which was considered appropriate for the age of students in the study. These items yielded a Self-Efficacy Scale and provided a total Self-Efficacy Score ranging from 5 to 15. A mean score was calculated for each student for this scale, given that a mean score indicates the student's average perception across the items.

Independent Variables

The independent variables were the following mathematics interventions: (a) repeated practice (RP), (b) folding-in technique (FI), and (c) immediate corrective feedback (ICF). Each of these treatment conditions is described below.

Treatment Conditions

RP. During RP, the student completed an instructional worksheet three times, untimed, before DCPM and EPM were assessed (e.g., Cates, 2005). Then, the student completed an intervention monitoring worksheet that included the same problems as in the corresponding instructional worksheet, but with a different order to avoid memorization of the answers. The student was allowed 1 minute to complete the intervention monitoring worksheet. The interventionist calculated DCPM and EPM for the intervention monitoring worksheet. No corrective feedback was provided to the student during RP.

FI. During FI, the interventionist, *prior to intervention,* assessed known and unknown facts for each student using approximately 100 instructional level mathematics flashcards (e.g., one-digit plus one-digit addition sums to 18) (e.g., Cooke et al., 1993; Cooke & Reichard, 1996). Specifically, the interventionist presented each of the flashcards to the student, and the fact was considered as known if he or she responded correctly within 3 seconds and as unknown otherwise. In each intervention session, an interventionist presented 10 mathematics fact flashcards including 7 (70%) known facts and 3 (30%) unknown facts, and a student responded to each fact. If the student made an error, the interventionist immediately corrected the student without teaching the calculation procedure. The presentation of the same 10 flashcards was repeated three times with a different order (i.e., cards were randomly shuffled) each time. Finally, the student completed an intervention monitoring worksheet that included the same problems as those used during FI. The student was allowed 1 minute to complete the worksheet. The interventionist calculated DCPM and EPM for the intervention monitoring worksheet. It should be noted that during FI, intervention monitoring worksheets also included 70% of known facts and 30% of unknown facts to promote students' acquisition of both types of facts.

After each session, if the student correctly responded to the same known fact on an intervention monitoring worksheet in three consecutive sessions, this fact was considered to be at mastery and removed from the known pile. If the student correctly responded to the same unknown fact on the worksheet in all three consecutive sessions, this fact was placed into the known pile. If the student made an error on the same known fact in all three consecutive sessions, this fact was considered to have become unknown and was placed in the unknown pile. If the student made an error on a fact in one session, which was not the three consecutive times, but responded to it correctly in the next session, the fact was kept in the same pile. For example, if a student responded to the same unknown fact correctly two consecutive times, but made an error the third time, this fact was remained in the unknown pile. Additionally, the fact was remained in the unknown pile, until the student responded to this fact correctly three more consecutive sessions. Facts that were not attempted on a worksheet were remained in the same pile. For example, facts that were not solved within 1 minute could not be determined either known or unknown facts. Thus, these facts were remained in the same pile. As flashcards in the unknown pile decreased (e.g., less than 10 unknown facts), evaluation of known and unknown facts was administered again using additional flashcards at the student's instructional level to determine unknown facts to be included in the procedure.

ICF. During ICF, the student completed each problem one by one on an instructional worksheet (Bennett & Cavanaugh, 1998; Skinner et al., 1992). Immediately after the student completed each problem, the interventionist checked the answer and told the student if it was correct or incorrect (e.g., "Right!" or "Wrong."). If the answer was correct, the interventionist had the student solve the next problem. If the answer was incorrect, the interventionist provided the student with the right answer (e.g., "You answered 3 + 5 = 7, but the answer is 8”), briefly explained the calculation procedure (e.g., “Here is how you do it; 3, then add 5 so 4, 5, 6, 7, and 8”), and let the student solve the next problem. Finally, the student completed an intervention monitoring worksheet including the same problems as those covered during the instructional worksheet, provided in a random order. The student was allowed 1 minute to complete the intervention monitoring worksheet. The interventionist calculated DCPM and EPM for the intervention monitoring worksheet.

Dependent Variables

The dependent variables were (a) DCPM on the intervention monitoring worksheets, (b) DCPM on the maintenance worksheets, (c) DCPM on the generalization worksheets, (d) EPM on the intervention monitoring worksheets, (e) EPM on the maintenance worksheets, (f) EPM on the generalization worksheets, and (g) ratings on the social validity checklist (see the procedures section for details on maintenance and generalization assessment). DCPM was the sum of all digits answered correctly on a mathematics calculation worksheet within the 1 minute time limit. EPM was the number of errors on a mathematics worksheet completed within the 1 minute time limit. See Appendix J for a description on the calculation of DCPM and EPM.

Procedures

All the procedures were implemented in an empty classroom, the library, or a similar room at the students’ school. The room was quiet and had adequate lighting. The room was occupied by no more than two interventionists (the researcher and another interventionist) and the two students. The two interventionists sat separately to prevent distraction. Each session lasted for approximately 20 to 30 minutes. The following sections describe the procedures used in the study including: (a) training of an interventionist, (b) skill deficit assessment (i.e., determination that each participant meets selection criteria), (c) pre-treatment assessment, (d) experimental analysis (brief experimental analysis and extended analysis), (e) baseline, (f) intervention implementation, (g) maintenance assessment, (h) generalization assessment, (i) social validity assessment, and (j) follow-up assessment. Procedures of assessment for interscorer agreement and treatment integrity are also described.

Training of the Interventionist

A school psychology graduate student assisted the researcher in conducting the study as an interventionist. The interventionist was trained in the procedures prior to the study. The researcher trained the interventionist on how to administer CBM and implement the interventions utilized in the study. Regarding CBM, the interventionist mastered how to determine students' frustrational, instructional, and mastery levels (Deno & Mirkin, 1977, cited in Shapiro & Lentz, 1986; see Appendix B) and how to calculate DCPM and EPM using a sample mathematics worksheet (see Appendix J). Then, paired, the interventionist administered a worksheet with the researcher for 1 minute and practiced scoring the worksheet and determining a level (see Appendix K). Integrity was trained to 100% criteria.

Next, the interventionist mastered how to implement the three interventions utilized in the study. The researcher provided a treatment integrity checklist for each intervention on which a step-by-step procedure was listed (see Appendix L to N), modeled each intervention, had the interventionist practice implementing each intervention, and provided immediate corrective feedback regarding performance of the interventionist. The researcher then checked the treatment integrity of the interventionist for intervention implementation. The interventionist demonstrated 100% integrity for each intervention. Finally, the interventionist mastered the follow-up procedure. Integrity was measured using an integrity checklist (see Appendix O) with 100% integrity for the follow-up procedure.

Pre-Treatment Assessment

Pre-treatment assessment was conducted to determine the current grade level of each student using CBA. Specifically, each student completed three sets of mixed-skill mathematics worksheets at the student's grade, median DCPM and EPM were calculated, and the level of the student at his or her grade was determined. For a student whose level was frustrational, three sets of mathematics worksheets one grade lower were administered to determine a level at the grade. This procedure was continued until the student's both instructional and mastery levels were identified. Finally, three sets of single-skill worksheets at the student's instructional level were administered. Students whose grade levels were frustrational on the first-grade materials (e.g., students who have difficulties on number sense or counting) were excluded from this study.

Skill Deficit Assessment

Skill deficit assessment was conducted for all students, using CBM, to determine adherence to selection criteria for the study (i.e., had a skill deficit rather than a performance deficit in mathematics calculation). The following procedure was used in the skill deficit assessment. First, the student was told his or her median score for the instructional grade identified in the CBA. Then, the student was told that a large reward (e.g., two favorite things from a treasure box such as candies, gums, and toys) would be provided if he or she could surpass that score by 50% or more (Duhon et al., 2004) but a small reward (e.g., one favorite thing from a treasure box) would be provided otherwise. One more CBA worksheet at the instructional grade was administered, DCPM was calculated, and a large or small reward was provided to the student depending on the DCPM. Students who did not surpass the median score by 50% were identified as having skill deficit and were included in this study. Conversely, those who surpassed the median score by 50% or above were considered to exhibit a performance deficit and, thus, were excluded from the study. One general education student was excluded from this study following determination of a performance deficit.

Experimental Analysis

An experimental analysis was conducted to identify the most effective intervention for each student. The experimental analysis consisted of the following two analyses: (a) brief experimental analysis (BEA) and (b) extended analysis.

BEA. During BEA, the following series of analyses were conducted for all students in general and special education. First, three interventions were implemented to determine the most effective intervention that would be used in the intervention phase following the BEA. These interventions included: (a) RP, (b) FI, and (c) ICF. The order of the interventions in the BEA was randomized for each student to prevent sequential effects (see Figure 2.1).

Student	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Jonathan	FI	RP	ICF	RP	RP	ICF	FI	ICF	FI	RP	FI	ICF	RP	–
Jamario	ICF	RP	FI	FI	ICF	FI	RP	FI	RP	ICF	RP	FI	ICF	–
Jacklyn	RP	FI	ICF	ICF	FI	ICF	RP	FI	RP	ICF	FI	ICF	FI	RP
Kathy	RP	FI	ICF	RP	RP	ICF	FI	RP	FI	ICF	FI	ICF	RP	–
Kaley	ICF	RP	FI	FI	–	–	–	–	–	–	–	–	–	–
Tyson	RP	ICF	FI	ICF	RP	FI	ICF	FI	ICF	RP	FI	RP	ICF	–

Figure 2.1 Intervention Order during Experimental Analysis

Note. RP = repeated practice; FI = folding-in technique;
ICF = immediate corrective feedback.

Then, the intervention(s) that produced the highest DCPM (e.g., DCPM which was 20% higher than the DCPM of the other two interventions) were replicated in the final session of the BEA to confirm its effects. It should be noted that interventions that increased students' performance by 20% above baseline were considered to be effective in previous studies involving BEA in reading interventions (e.g., Jones & Wickstrom, 2002; Noell, Freeland, Witt, & Gansle, 2001); however, this fact has not been examined in previous studies in BEA for mathematics.

Finally, if the student demonstrated DCPM 20% higher than the DCPM of the other two interventions twice under the same intervention during BEA (i.e., initial session and replicated session for the same intervention), the intervention was considered differentially effective for the student. Based on this criterion (i.e., one intervention was considered differentially effective over the other two interventions), the student was included in this study. Only Kaley met this criterion, who demonstrated divergence in DCPM between FI and the other two interventions in BEA (FI was selected for Kaley); thus, an extended analysis was not conducted.

Extended analysis. If a student demonstrated DCPM that was 20% higher in one intervention than in the other interventions once, but did not replicate in the BEA, or if a student did not demonstrate DCPM that was 20% higher than that of the other interventions under any interventions, an extended analysis was then conducted. In the extended analysis, an alternating treatments design (ATD; Hayes, Barlow, & Nelson-Gray, 1999) was conducted after the replicated session to further examine the effects of the three interventions (i.e., RP, FI, and ICF). Specifically, the three interventions were implemented with a randomized order for nine sessions (i.e., three sessions for each intervention) using an ATD, and their effects were visually analyzed in terms of divergence for intervention effects (Hayes et al.). For each student, if one intervention was more effective than the other two interventions based on their mean DCPM across the initial BEA and extended analysis (i.e., one intervention produced higher mean DCPM than that of the other two interventions), the intervention was considered differentially effective for the student, and thus he or she was included in this study. An extended analysis was conducted with all students except for Kaley. Overall, a total of 3 students in general education and 3 students in special education were included in this study. No further participant criterion analysis was conducted.

Pairing students with and without SLD. After the experimental analysis, a general education at-risk student and a special education student who had SLD in mathematics calculation, who responded to the same intervention, were matched in order to compare the intervention responsiveness of each type of student (i.e., general and special education). Three matched pairs of general and special education students, each of which responded to each intervention, were included in this study. The identified pairs and an intervention for each pair were as follows: (a) Jonathan and Kathy for RP; (b) Jamario and Kaley for FI; and (c) Jacklyn and Tyson for ICF. All students had a matched student based on selection criteria for at-risk general education students and special education students with SLD and response to intervention criteria.

Baseline

During each baseline session, one instructional worksheet at the student's instructional level was administered. A single skill was targeted. During baseline, if the data of one of the students in a pair demonstrated a continuing increasing trend, data collection was continued for both paired students until a stable level or a decreasing trend was observed. No intervention was implemented during baseline sessions. For two pairs of students, an increasing trend was found for four sessions; therefore, baseline included five sessions for these pairs.

Intervention Implementation

For each student, the most effective intervention identified in a BEA was implemented for 18 sessions (2 to 4 sessions per week for 6 weeks). Intervention progress was examined using intervention monitoring worksheets (2 to 4 days a week). A single skill at the student's instructional level was targeted during the intervention. Students in the same pair received the same intervention on the same day or on two consecutive days (e.g., one day for a student without SLD and the next day for a student with SLD in the pair); on a rare occasion, an intervention session was conducted with a pair of students across 3 to 4 days due to unforeseen events (e.g., absence for doctor appointment, family crisis, etc.) and these occurrences are stated in the results section. The intervention procedure was the same for each intervention that was implemented under BEA.

Maintenance Assessment

Maintenance of the performance at a student's instructional level was assessed using a maintenance worksheet (i.e., a cold worksheet). A single skill was targeted during the maintenance assessment. In each intervention session, prior to implementation of an intervention, a maintenance worksheet was administered for 1 minute. Following the administration, the interventionist calculated DCPM and EPM. No feedback was provided regarding the student's performance on the maintenance worksheet.

Generalization Assessment

Generalization from a student's instructional level to his or her current grade level was also assessed. Multiple skills were targeted during the generalization assessment. Every third session, after completing an intervention monitoring worksheet, a student completed a generalization worksheet for 1 minute. The interventionist calculated DCPM and EPM. No feedback was provided regarding the student's performance on the generalization worksheet.

Social Validity Assessment

The *Student Social Validity Checklist* was administered to assess students' satisfaction with interventions and sense of self-efficacy toward mathematics. Each student completed the *Student Social Validity Checklist, Satisfaction Scale* (see Appendix F to H) during the last intervention session (i.e., the last intervention session). The satisfaction scale was administered immediately after completing a generalization worksheet. Each student also completed the *Student Social Validity Checklist, Self-Efficacy Scale* (see Appendix I) during the pre-treatment assessment (before conducting CBA), the first intervention session (before completing a maintenance worksheet), and the last intervention session (after completing a generalization worksheet). For each student, a mean rating was calculated for the Satisfaction and Self-Efficacy Scale, respectively. For the Self-Efficacy Scale, mean rating change was also analyzed across the three assessment sessions.

Follow-up Assessment

For each student, follow-up data were collected at 1 week (7 days for all students except for Kathy; 8 days for Kathy since she was absent on the 7^{th} day due to a doctor's appointment) after the intervention was terminated. During the follow-up session, a student completed three instructional worksheets and one generalization worksheet, respectively. A single skill at the student's instructional level was targeted on the instructional worksheets, whereas multiple skills at the student's current grade were targeted on the generalization worksheet. An interventionist calculated the median of DCPM and EPM for the instructional worksheets and DCPM and EPM for the generalization worksheet. No intervention was implemented.

Interscorer Agreement

Interscorer agreement (ISA) for DCPM and EPM was calculated for worksheets used in the pre-treatment assessment, BEA, intervention, and follow-up sessions. For each student, the researcher randomly selected 34.3% of all worksheets which were equally distributed across all phases, and a second trained graduate student re-scored all answers on the worksheets and re-calculated DCPM and EPM for these worksheets. ISA was calculated using the following formula:

$$\text{ISA} = \frac{\text{Number of Agreements}}{\text{Number of Agreement} + \text{Number of Disagreement}} \times 100 \qquad (2\text{-}1)$$

The mean ISA was 99.3% (range = 75-100); miscalculation of DCPM was observed on some worksheets (e.g., 1 DCPM for an addition problem with a two-digit answer such as 8+2), leading to less than perfect scoring.

ISA was also calculated for randomly-selected 33.3% of all satisfaction and self-efficacy checklists which were equally distributed across the study (i.e., only post-study assessment for satisfaction scales; pre-, mid- and post-study for self-efficacy scales). The trained second scorer re-scored these checklists and ISA was calculated for each checklist using the same formula stated above. ISA was 100% for all checklists.

Treatment and Procedural Integrity

Treatment integrity is defined as the degree to which a treatment (intervention) is implemented as intended (Yeaton & Sechrest, 1981). Researchers have stated the necessity and importance of assessment of treatment integrity in the implementation of research (Gresham, 1989; Gresham, Gansle, & Noell, 1993; Gresham, Gansle, Noell, Cohen, & Rosenblum, 1993). Without treatment integrity, it is questionable if the behavior change that occurred is due to the treatment or to some extraneous variable (Sterling-Turner, Watson, & Moore, 2002; Sterling-Turner, Watson, Wildmon, Watkins, & Little, 2001).

In this study, treatment (e.g., interventions) or procedural (e.g., CBA) integrity was assessed using an integrity checklist, on which all procedural steps were listed, for CBA, each intervention, and follow-up assessment (See Appendix K to O). For each student, treatment or procedural integrity was checked for 100% of the sessions across all phases. Specifically, immediately after each of the sessions, an interventionist completed an integrity checklist by marking each item on it. Additionally, 6.3% of all sessions, equally distributed across all phases except for the follow-up session, were observed by a trained observer to confirm integrity. Integrity was calculated using the following formula:

$$\text{Integrity} = \frac{\text{Number of Completed Steps}}{\text{Number of Total Steps}} \times 100 \qquad (2\text{-}2)$$

Integrity was reported as 100% for all sessions. Agreement for integrity was 100% as assessed by the trained observer.

Design and Analysis

For each student, a BEA was used to identify the most effective intervention. Evaluation of the intervention was conducted using a simple phase change (i.e., A+B with A indicating baseline and B indicating an intervention phase) with a follow-up session. Specifically, after pre-treatment assessment, the effects of the three interventions were compared using a BEA for each student. The order of the interventions was counterbalanced to prevent sequential effects; that is, the interventions were not implemented in the same order for any more than two students in general or special education (see Table 2.4 for intervention order). An ATD was used if a clear differentiation in the effects of the three interventions was not observed in the first four sessions during BEA, in which the three interventions were implemented in random alternating order for nine sessions (three sessions for each intervention) and the most effective intervention was identified through visual inspection of data points (i.e., meeting differentiation criteria between one intervention and the other two interventions). For each student, the most effective intervention identified in the BEA was implemented in the intervention phase. The intervention was implemented for 18 sessions. Follow-up data were collected at 1 week after termination of the intervention.

Hypothesis One

Hypothesis One stated that students, with and without SLD in mathematics calculation, would demonstrate responsiveness when an intervention was identified through an experimental analysis and implemented for approximately 6 weeks. To examine Hypothesis One, data were analyzed through visual inspection of level, trend, and variability for each student. Additionally, the following two data were calculated for each student to evaluate his or her responsiveness to intervention: (a) a median or mean DCPM and a median or mean EPM in the pre-treatment, BEA, intervention, and follow-up sessions and (b) an individually-derived estimated weekly slope from pre-treatment assessment to an intervention phase, calculated by subtracting a median pre-treatment DCPM from a mean intervention DCPM and dividing by the total number of weeks spent for pre-treatment, experimental analysis, and intervention sessions (Shapiro, 1996, 2004). These data were compared between paired students (general education academically at-risk students and special education students with SLD in mathematics calculation) descriptively rather than statistically. Additionally, students whose weekly slopes from pre-treatment assessment to an intervention phase were at or above .50 DCPM for the third grade and at or above 1.20 DCPM for the fourth grade, suggested by Fuchs et al. (1993) as an aspirational goal for each of these grades, were considered to be responders. The number of responders was calculated and reported for general education academically at-risk students and special education students with SLD in mathematics calculation.

Hypothesis Two

Hypothesis Two stated that students, with and without SLD in mathematics calculation, would demonstrate satisfaction and self-efficacy following implementation of the intervention that best addressed their skill deficit. To examine Hypothesis Two, a mean rating of the *Student Social Validity Checklist, Satisfaction Scale* administered during the last intervention session (i.e., the last intervention session) and that of the *Student Social Validity Checklist, Self-Efficacy Scale* administered during the pre-treatment assessment (before conducting CBA), the first intervention session (at the beginning of the session), and the last intervention session (at the end of the session) were calculated for the students. For satisfaction and self-efficacy, the mean ratings of the two groups were compared descriptively rather than statistically due to a small number of participants in the two groups. For self-efficacy, mean rating change was also examined across the three assessment sessions.

CHAPTER III

RESULTS

In this chapter, the results of the experimental analysis (brief experimental analysis and extended analysis), baseline, interventions, maintenance, generalization, and follow-up are stated. The following data for these phases were analyzed: (a) digits correct per minute (DCPM) on the intervention monitoring worksheets, (b) DCPM on the maintenance worksheets, (c) DCPM on the generalization worksheets, (d) errors per minute (EPM) on the intervention monitoring worksheets, (e) EPM on the maintenance worksheets, (f) EPM on the generalization worksheets. Results of social validity (satisfaction and self-efficacy) are also stated. These analyses were conducted to examine the two hypotheses stated in Chapter I.

Experimental Analysis

After students who met the selection criteria of this study were identified based on the curriculum-based assessment (CBA) and skill-deficit assessment, an experimental analysis was conducted to identify the most effective intervention for each student (See Table 2.2 for intervention order for each student). In the experimental analysis, a brief experimental analysis (BEA) was conducted first with each student to examine whether one intervention was differentially effective than the others for the student (e.g., an intervention that produced DCPM that was 20% higher than that of the other interventions). During the BEA, the following three interventions were implemented: (a) repeated practice (RP), (b) folding-in technique (FI), and (c) immediate corrective feedback (ICF). For each student, the intervention that produced the highest DCPM during the three interventions was also replicated with the student to confirm its effects. For students who did not demonstrate DCPM that was 20% higher under the same intervention twice than that of the other interventions, an extended analysis was also conducted, in which the three interventions were implemented in a randomized order for nine sessions (i.e., three sessions for each intervention). Results of an experimental analysis for each student are summarized below. Table 3.1 provides data on instructional worksheets in the CBA and data on intervention monitoring worksheets during the experimental analysis for each student.

Experimental Analysis Results for the Students in General Education

Experimental analysis results for the academically at-risk students in general education are presented below. The most effective intervention for each of the students identified during the experimental analysis is also stated. CBA and experimental analysis results for pairs of students in general and special education are also summarized on Table 3.1. Figures 3.1 to 3.3 display DCPM and EPM under the CBA and experimental analysis for the students in general education.

Table 3.1 Median Digits Correct Per Minute (DCPM) during CBA and Mean and Median DCPM during Experimental Analysis

	Placement	CBA	Experimental Analysis					
			RP		FI		ICF	
Student		*Mdn*	*M*	*Mdn*	*M*	*Mdn*	*M*	*Mdn*
			Dyad One					
Jonathan	General	10.0	**19.8** (4.4)	20.0	16.8 (3.8)	16.0	17.8 (3.2)	17.5
Kathy	SPED	19.0	**26.8** (2.6)	28.0	23.8 (1.3)	24.0	26.3 (3.1)	27.0
			Dyad Two					
Jamario	General	15.0	15.3 (1.0)	15.5	**17.6** (2.1)	18.0	16.3 (4.8)	17.5
Kaley	SPED	16.0	12.0 (N/A)	N/A	**20.5** (N/A)	N/A	14.0 (N/A)	N/A
			Dyad Three					
Jacklyn	General	13.0	18.5 (3.1)	17.5	17.8 (1.8)	18.0	**21.0** (3.3)	20.0
Tyson	SPED	32.0	38.8 (2.6)	39.5	34.5 (6.2)	34.5	**40.8** (4.1)	41.0

Note. CBA = Curriculum-based assessment; RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback; *M* = mean; *Mdn* = median; General = general education; SPED = special education; bold indicates highest DCPM for each student; data in parentheses indicate standard deviations.

Jonathan. Figure 3.1 displays Jonathan's calculation performance on instructional worksheets during CBA and while completing intervention monitoring worksheets during experimental analysis. Jonathan obtained 10 DCPM and 1 EPM for the first grade during the CBA. During the BEA, he obtained 16, 20, and 15 DCPM under the FI, RP, and ICF intervention conditions, respectively. He also obtained 0 EPM under each of the three intervention conditions. Jonathan demonstrated a 20% increase in DCPM under the RP intervention condition once. Because he demonstrated the highest DCPM under the RP intervention condition, the RP intervention was replicated. Jonathan obtained 18 DCPM and 0 EPM during the replicated RP session. Given that he did not demonstrate a 20% increase twice under the RP intervention condition, an extended analysis was conducted with him.

Across the experimental analysis sessions, Jonathan obtained a mean of 19.8 DCPM (range = 14-26), 16.8 DCPM (range = 13-22), and 17.8 DCPM (range = 15-21) under the RP, FI, and ICF intervention conditions, respectively. He also obtained a mean of 0 EPM under the three intervention conditions. Jonathan increased DCPM level from CBA to the experimental analysis under all intervention conditions. Divergence was observed under the three intervention conditions with the RP intervention shown to be superior to the other two intervention conditions. Jonathan maintained low EPM under all intervention conditions. Overall, the RP intervention was identified as the most effective intervention for Jonathan.

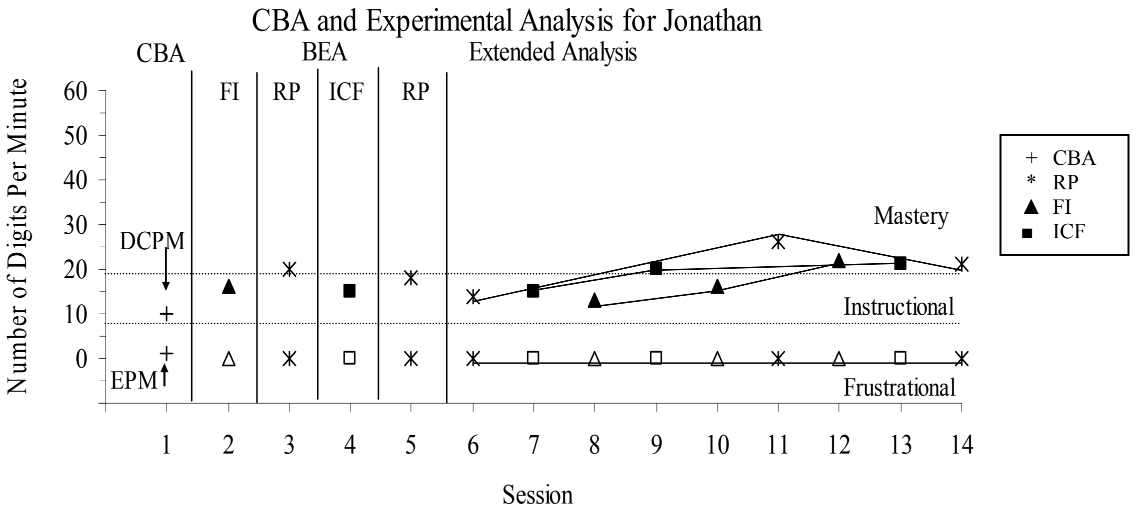

Figure 3.1 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA), Brief Experimental Analysis (BEA), and Extended Analysis for Jonathan

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Jamario. Figure 3.2 displays Jamario's calculation performance on instructional worksheets during CBA and while completing intervention monitoring worksheets during experimental analysis. Jamario obtained 15 DCPM and 0 DCPM for the first benchmark of the second grade (i.e., a skill for one-digit plus one-digit addition sums to 18) during the CBA. During the BEA, he obtained 15, 15, and 20 DCPM under the ICF, RP, and FI intervention conditions, respectively. He also obtained 0, 1, and 0 EPM under each of the three intervention conditions, respectively. Jamario demonstrated a 20% increase under the FI intervention condition once. Because he demonstrated the highest DCPM under the FI intervention condition, the FI intervention was replicated. He obtained 16 DCPM and 0 EPM during the replicated FI session. Given that he did not demonstrate a 20%

increase twice under the FI intervention condition, an extended analysis was conducted with him.

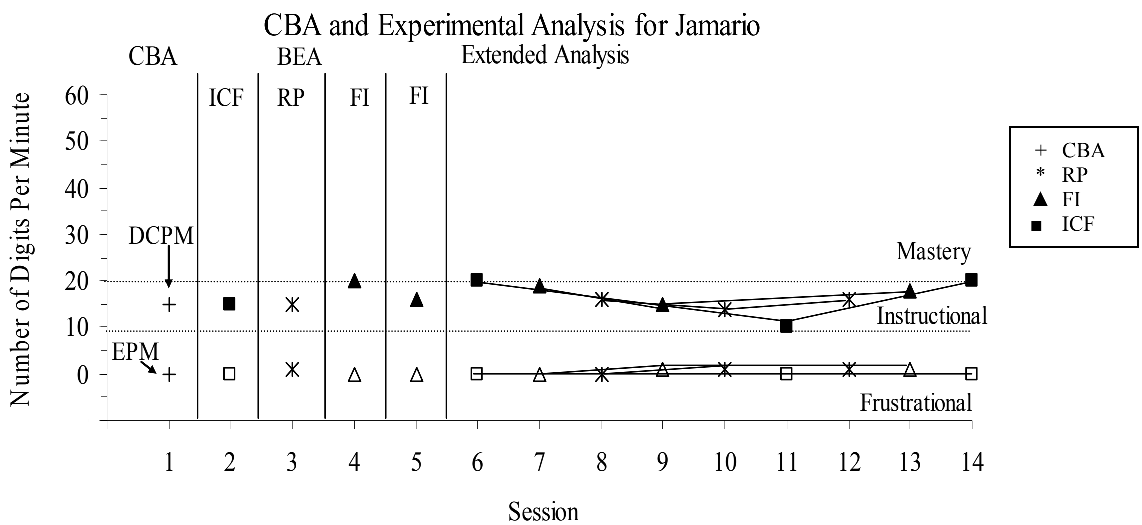

Figure 3.2 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA), Brief Experimental Analysis (BEA), and Extended Analysis for Jamario

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Across the experimental analysis sessions, Jamario obtained a mean of 15.3 DCPM (range = 14-16), 17.6 DCPM (range = 15-20), and 16.3 DCPM (range = 10-20) under the RP, FI, and ICF intervention conditions, respectively. He also obtained a mean of 0.8 EPM (range = 0-1), 0.4 EPM (range = 0-1), and 0 EPM under each of the three intervention conditions, respectively. Jamario increased DCPM level from CBA to the experimental analysis under all intervention conditions. Divergence was observed under the three intervention conditions with the FI intervention shown to be superior to the other two intervention conditions. Jamario maintained low EPM under all intervention conditions. Overall, the FI intervention was identified as the most effective intervention for Jamario.

Jacklyn. Figure 3.3 displays Jacklyn's calculation performance on instructional worksheets during CBA and while completing intervention monitoring worksheets during experimental analysis. Jacklyn obtained 13 DCPM and 0 EPM for the first grade during the CBA. During the BEA, she obtained 17, 20, and 20 DCPM under the RP, FI, and ICF intervention conditions, respectively. She also obtained 1, 0, and 0 EPM under each of the three intervention conditions, respectively. Jacklyn did not demonstrate a 20% increase under any intervention conditions. Because Jacklyn demonstrated the highest DCPM under the FI and ICF intervention conditions, these two interventions were replicated in reverse order. In the replicated sessions, she obtained 20 and 16 DCPM under the ICF and FI intervention conditions, respectively. She also obtained 0 EPM under each of the two intervention conditions. Because Jacklyn did not demonstrate a 20% increase twice under any intervention conditions during the BEA, an extended analysis was conducted with her.

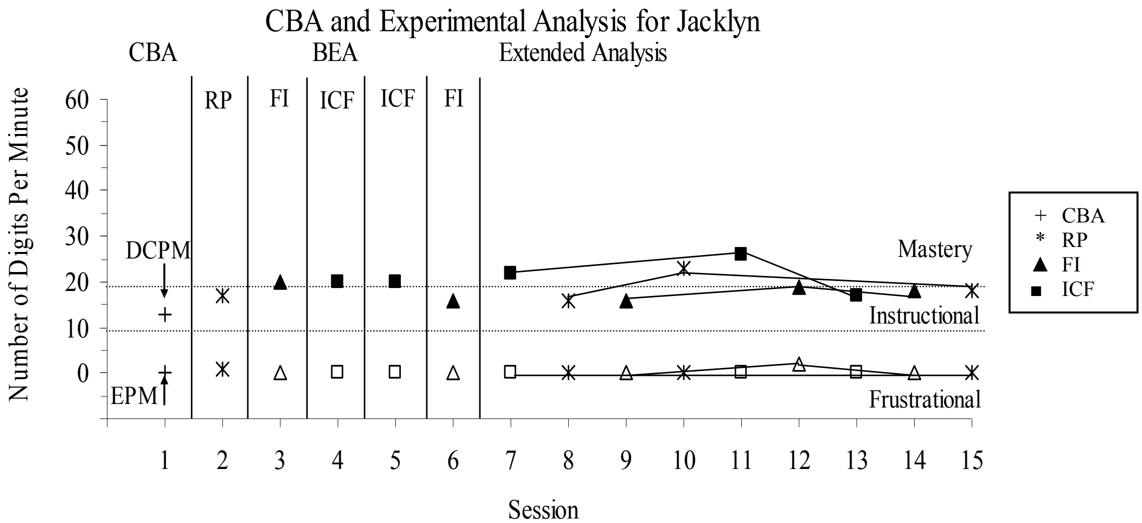

Figure 3.3 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA), Brief Experimental Analysis (BEA), and Extended Analysis for Jacklyn

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Across the experimental analysis sessions, Jacklyn obtained a mean of 18.5 DCPM (range = 16-23), 17.8 DCPM (range = 16-20), and 21 DCPM (range = 17-26) under the RP, FI, and ICF intervention conditions, respectively. She also obtained a mean of 0.3 EPM (range = 0-1), 0.4 EPM (range = 0-2), and 0 EPM under each of the three intervention conditions, respectively. Jacklyn increased DCPM level from CBA to the experimental analysis under all intervention conditions. Divergence was observed in the three interventions with the ICF intervention shown to be superior to the other two intervention conditions. Jacklyn maintained low EPM under all intervention conditions. Overall, the ICF intervention was identified as the most effective intervention for Jacklyn.

Experimental Analysis Results for the Students in Special Education

Experimental analysis results for students in special education are presented in this section. The most effective intervention identified for each of the students is also stated. Figures 3.4 to 3.6 display DCPM and EPM under the CBA and experimental analysis for the students in special education.

Kathy. Figure 3.4 displays Kathy's calculation performance on instructional worksheets during CBA and while completing intervention monitoring worksheets during experimental analysis. Kathy obtained 19 DCPM and 3 EPM for the first grade during the CBA. During the BEA, she obtained 24, 22, and 22 DCPM under the RR, FI, and ICF intervention conditions, respectively. She also obtained 1, 0, and 2 EPM under each of the three intervention conditions, respectively. Kathy did not demonstrate a 20% increase in DCPM under any interventions. Because she demonstrated the highest DCPM under the RP intervention condition, the RP intervention was replicated. She obtained 24 DCPM and 2 EPM during the replicated RP session. Given that she did not demonstrate a 20% increase twice under any intervention conditions, an extended analysis was conducted with her.

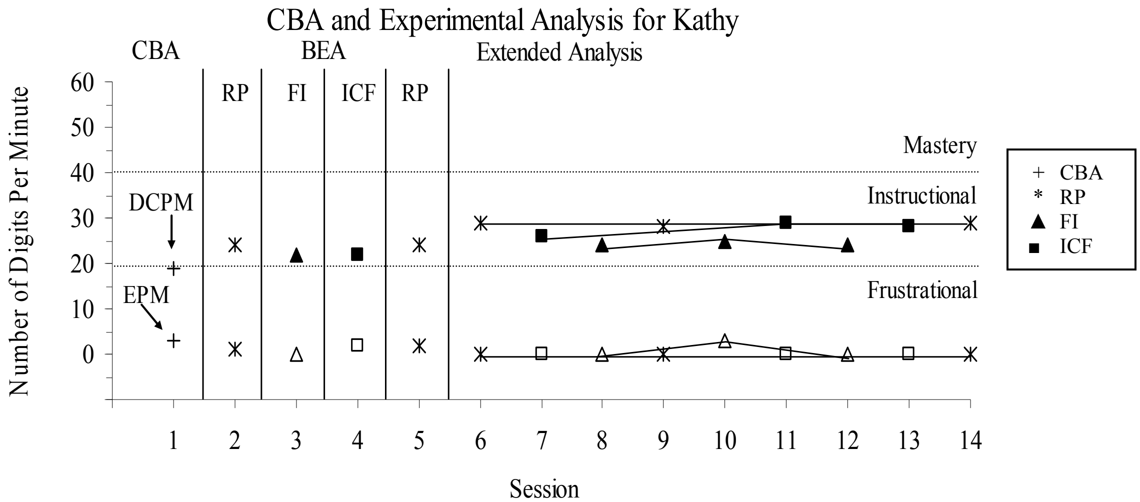

Figure 3.4 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA), Brief Experimental Analysis (BEA), and Extended Analysis for Kathy

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Across the experimental analysis sessions, Kathy obtained a mean of 26.8 DCPM (range = 24-29), 23.8 DCPM (range = 22-25), and 26.3 DCPM (range = 22-29) under the RP, FI, and ICF intervention conditions, respectively. She also obtained a mean of 0.6 EPM (range = 0-2), 0.8 EPM (range = 0-3), and 0.5 EPM (range = 0-2) under each of the three intervention conditions, respectively. Kathy increased DCPM level from CBA to the experimental analysis under all intervention conditions. Divergence was observed between the FI intervention and the other two interventions, but not between the RP and ICF interventions. Kathy maintained low EPM under all intervention conditions. During the experimental analysis, the RP and ICF interventions produced higher mean DCPM than the FI intervention. A difference in DCPM between the RR and ICF interventions

(0.5 DCPM difference) was also negligible. However, Kathy frequently used primitive strategies (e.g., finger counting) on basic addition facts for the first grade level, indicating her lack of automaticity. Given that a repeated practice component has been stated as effective and crucial in building automaticity in students with calculation deficits (e.g., Cates, 2005; Hasselbring et al., 1988), the RP intervention was selected for Kathy.

Kaley. Figure 3.5 displays Kaley's calculation performance on instructional worksheets in CBA and while completing intervention monitoring worksheets during experimental analysis. Kaley obtained 16 DCPM and 1 EPM for the first grade during the CBA. During the BEA, she obtained 14, 12, and 21 DCPM under the ICF, RP, and FI intervention conditions, respectively. She also obtained 2, 3, and 0 EPM under each of the three intervention conditions, respectively. Kaley demonstrated a 20% increase under the FI intervention. Because she demonstrated the highest DCPM under the FI intervention condition, the FI intervention was replicated. Kaley obtained 20 DCPM and 0 EPM during the replicated FI session. Because Kaley demonstrated a 20% increase in DCPM twice under the FI intervention condition, an extended analysis was not conducted with her. Kaley increased DCPM level from CBA to BEA under the FI intervention, whereas it decreased under the RP and ICF interventions. Kaley also maintained low EPM under all intervention conditions. Overall, the FI intervention was identified as the most effective intervention for Kaley.

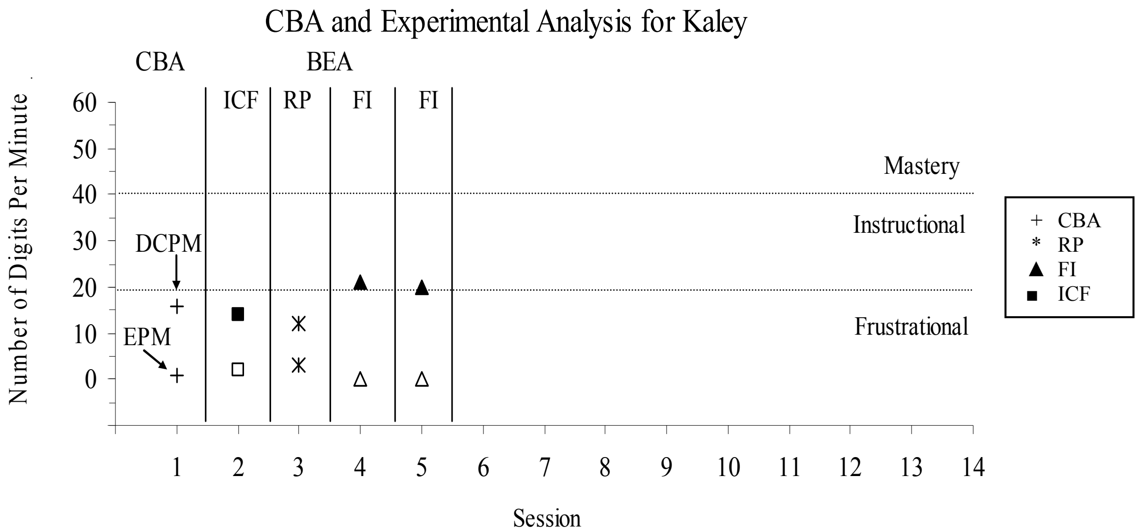

Figure 3.5 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA) and Brief Experimental Analysis (BEA) for Kaley

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Tyson. Figure 3.6 displays Tyson's calculation performance on instructional worksheets during CBA and while completing intervention monitoring worksheets during experimental analysis. Tyson obtained 32 DCPM and 0 EPM for the first benchmark of the second grade during the CBA. During the BEA, Tyson obtained a mean of 35, 41, and 36 DCPM under the RR, ICF, and FI intervention conditions, respectively. He also obtained a mean of 0, 0, and 1 EPM under each of the three interventions, respectively. Tyson did not demonstrate a 20% increase under any interventions. Because he demonstrated the highest DCPM under the ICF intervention condition, the ICF intervention was replicated. He obtained 35 DCPM and 1 EPM during the replicated ICF session. Because Tyson did not demonstrate a 20% increase in DCPM twice under any intervention conditions, an extended analysis was conducted with him.

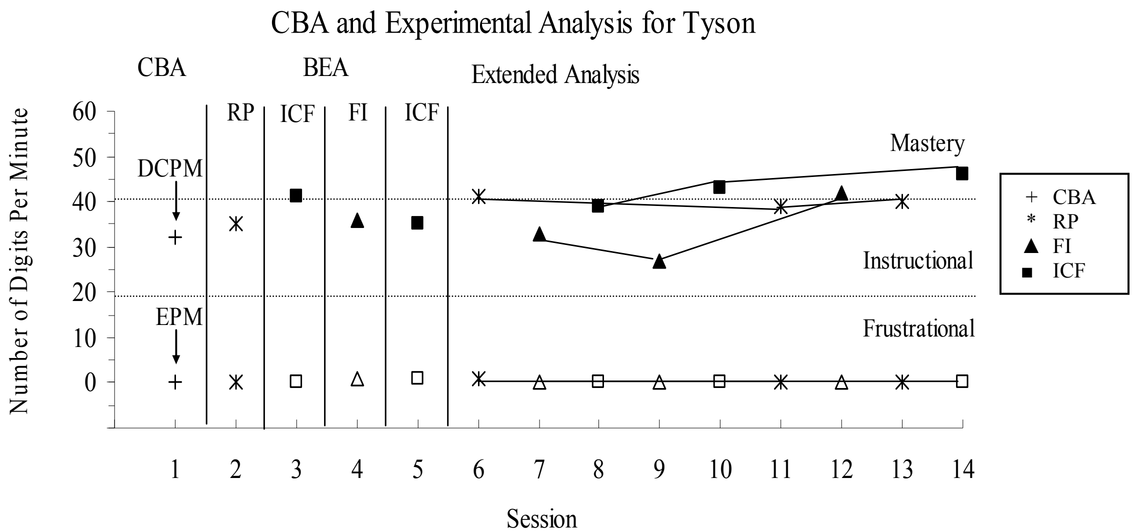

Figure 3.6 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) during Curriculum-Based Assessment (CBA), Brief Experimental Analysis (BEA), and Extended Analysis for Tyson

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback.

Across the experimental analysis sessions, Tyson obtained a mean of 38.8 DCPM (range = 35-41), 34.5 DCPM (range = 27-42), and 40.8 DCPM (range = 35-46) under the RP, FI, and ICF intervention conditions, respectively. He also obtained a mean of 0.3 EPM (range = 0-1), 0.3 EPM (range = 0-1), and 0.2 EPM (range = 0-1) under the RP, FI, and ICF intervention conditions, respectively. Tyson increased DCPM level from CBA to the experimental analysis under all intervention conditions. Divergence in DCPM was observed under the three intervention conditions with the ICF intervention shown to be superior to the other two intervention conditions. He maintained low EPM under all intervention conditions. Overall, the ICF intervention was identified as the most effective intervention for Tyson.

Effects of Empirically-Derived Interventions

Hypothesis One

Students, with and without SLD in mathematics calculation, would demonstrate responsiveness when an intervention was identified through a brief experimental analysis and implemented for approximately 6 weeks.

In this section, data regarding Hypothesis One is presented. After the most effective intervention was identified through the experimental analysis, baseline was introduced followed by an intervention phase including 18 sessions. Prior to baseline, students with and without SLD, who responded to the same intervention during the experimental analysis, were paired to compare their responsiveness. Maintenance (i.e., measurement of skill when an intervention has not been implemented for a period of time) data were collected prior to the implementation of the intervention every session. Generalization (i.e., measurement of skill generalized from instructional-level materials to grade-level materials) data were also collected every three intervention sessions. Follow-up data were collected on instructional and generalization worksheets one week after the termination of the intervention. The data of each student are described below. Tables 3.2 and 3.3 provides data on instructional, intervention monitoring, and maintenance worksheets for general and special education students, respectively; Table 3.4 provides data on generalization worksheets; Table 3.5 provides weekly slopes and decision of responders and nonresponders.

Table 3.2 Mean and Median DCPM and Range on Instructional, Intervention Monitoring, and Maintenance Worksheets during Baseline, Intervention, and Follow-Up for the Students in General Education

		Baseline			Intervention			F
Intervention	WS	*M*	*Mdn*	*Range*	*M*	*Mdn*	*Range*	*Mdn*
				Jonathan				
RP	I/IM	21.3 (2.9)	23.0	18.0 - 23.0	20.2 (4.6)	19.5	14.0 - 30.0	22.0
	MA				20.6 (5.3)	23.0	11.0 - 29.0	
				Jamario				
FI	I/IM	20.4 (4.4)	21.0	14.0 - 26.0	22.9 (3.8)	23.0	16.0 - 29.0	32.0
	MA				21.8 (3.5)	21.0	17.0 - 28.0	
				Jacklyn				
ICF	I/IM	20.2 (4.4)	21.0	13.0 - 25.0	22.9 (3.8)	24.0	13.0 - 28.0	28.0
	MA				23.8 (3.1)	23.0	18.0 - 30.0	

Note. RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback; WS = worksheet; I/IM = instructional/intervention monitoring worksheet; MA = maintenance worksheet; *M* = mean; *Mdn* = median; F = follow-up; data in parentheses indicate standard deviations.

Table 3.3 Mean and Median DCPM and Range on Instructional, Intervention Monitoring, and Maintenance Worksheets during Baseline, Intervention, and Follow-Up for the Students in Special Education

		Baseline			Intervention			F
Intervention	WS	*M*	*Mdn*	*Range*	*M*	*Mdn*	*Range*	*Mdn*
				Kathy				
RP	I/IM	24.0 (1.7)	23.0	23.0 - 26.0	29.7 (3.5)	30.0	24.0 - 35.4	28.0
	MA				28.8 (3.3)	29.0	22.0 – 33.0	
				Kaley				
FI	I/IM	23.2 (3.9)	25.0	19.0 - 27.0	19.8 (2.9)	20.5	15.0 - 26.0	22.0
	MA				20.3 (4.4)	20.0	12.0 - 29.0	
				Tyson				
ICF	I/IM	41.0 (2.1)	40.0	38.0 - 42.0	41.8 (4.1)	43.0	33.0 - 46.9	41.0
	MA				41.0 (4.6)	41.5	32.0 – 51.6	

Note. RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback; WS = worksheet; I/IM = instructional/intervention monitoring worksheet; MA = maintenance worksheet; *M* = mean; *Mdn* = median; F = follow-up; data in parentheses indicate standard deviations.

Table 3.4 Mean and Median DCPM and Range on Generalization Worksheets during Intervention and Follow-Up

	Intervention			Follow-Up
Student	*M*	*Mdn*	*Range*	*Mdn*
		General Education		
Jonathan[a]	5.7 (2.5)	5.5	3.0 - 9.0	6.0
Jamario[b]	9.2 (1.5)	9.5	7.0 - 11.0	15.0
Jacklyn[c]	8.2 (1.2)	8.5	6.0 - 9.0	8.0
		Special Education		
Kathy[a]	15.7 (0.8)	15.5	11.0 - 20.0	13.0
Kaley[b]	5.3 (2.1)	5.0	3.0 - 9.0	7.0
Tyson[c]	15.3 (1.8)	14.5	14.0 - 18.0	14.0

Note. *M* = mean, *Mdn* = median; [a] = repeated practice; [b] = folding-in technique; [c] = immediate corrective feedback; data in parentheses indicate standard deviations.

Table 3.5 Level Increase, Weeks Spent from CBA through the Intervention Phase, Weekly Slope, and Decision of Responders or Nonresponders

Student	Intervention	CBA	Level increase[a]	Weeks[b]	Weekly slope[c]	Decision
			General Education			
Jonathan	RP	10.0	10.2	12.0	0.9	Responder
Jamario	FI	15.0	7.9	13.0	0.6	Responder
Jacklyn	ICF	32.0	9.9	13.0	0.8	Responder
			Special Education			
Kathy	RP	19.0	10.7	9.0	1.2	Responder
Kaley	FI	16.0	3.8	8.0	0.5	Nonresponder
Tyson	ICF	32.0	9.8	12.0	0.8	Nonresponder

Note. A weekly slope was calculated by dividing level increase by the number of weeks spent from curriculum-based assessment (CBA) to the end of the intervention phase (i.e., $^{c} = {}^{a}/{}^{b}$). RP = repeated practice; FI = folding-in technique; ICF = immediate corrective feedback; [a] = level increase; [b] = weeks spent from CBA through the intervention phase; [c] = weekly slope.

Intervention Responsiveness of the Students in General Education

Responsiveness to empirically-derived intervention for student in general education is outlined in this section. Baseline, maintenance, generalization, and follow-up results are also outlined. Decision of responders or nonresponders is outlined. Figure 3.7 to 3.9 display data on instructional, intervention monitoring, and maintenance worksheets for students in general education.

Jonathan. The RP intervention was selected for Jonathan based on the experimental analysis results. Figure 3.7 displays Jonathan's calculation performance on instructional worksheets during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the RP intervention condition. During baseline, Jonathan obtained a mean of 21.3 DCPM and 0 EPM. He reached mastery level with nearly flat trend and little variability during baseline. After the RP intervention was introduced, on intervention monitoring worksheets, he showed a slight decrease in his fluency (*M* = 20.2 DCPM, range = 14-30), but maintained low error rates (*M* = 0.3 EPM, range = 0-2). On maintenance worksheets, he also showed a slight decrease in his fluency (*M* = 20.6 DCPM, range = 11-29), but maintained low rates of errors (*M* = 0.2 EPM, range = 0-2). He reached mastery level on both types of worksheets. For DCPM, an increasing trend was observed on intervention monitoring worksheets, whereas nearly flat trend was observed on maintenance worksheets. High variability in DCPM was observed on both types of worksheets. On generalization worksheets, Jonathan performed within frustrational level (*M* = 5.7 DCPM, range = 3-9) with low error rates (*M* = 0.8 EPM, range = 0-2). Low level, decreasing trend, and little variability in DCPM were observed on generalization worksheets. EPM was stable with little variability on all types of worksheets. During the follow-up session, Jonathan maintained his fluency (*Mdn* = 22 DCPM) and low error rates (*Mdn* = 0 EPM) on instructional worksheets. He scored within frustrational level (6 DCPM) with low error rates (2 EPM) on the one administered generalization worksheet. Jonathan's individually-derived estimated weekly slope (a rate of DCPM increase for a week) was 0.9 DCPM (12 weeks from CBA through the intervention phase). As such, he met the criterion of the weekly slope for third graders (.50 DCPM) suggested by Fuchs et al. (1993) and, thus, he was considered to be a responder. Overall, the RP intervention was shown to be effective in enhancing Jonathan's instructional skill. He also maintained his skill during and after the intervention phase. However, the RP intervention was not shown to be effective in enhancing his grade-level skills. Jonathan maintained low error rates on instructional-level and grade-level worksheets under the RP intervention. It should be noted that an intervention session was conducted with Jonathan across three days once and across four days once due to the unforeseen events of his paired student (e.g., Kathy's absence for her doctor appointment).

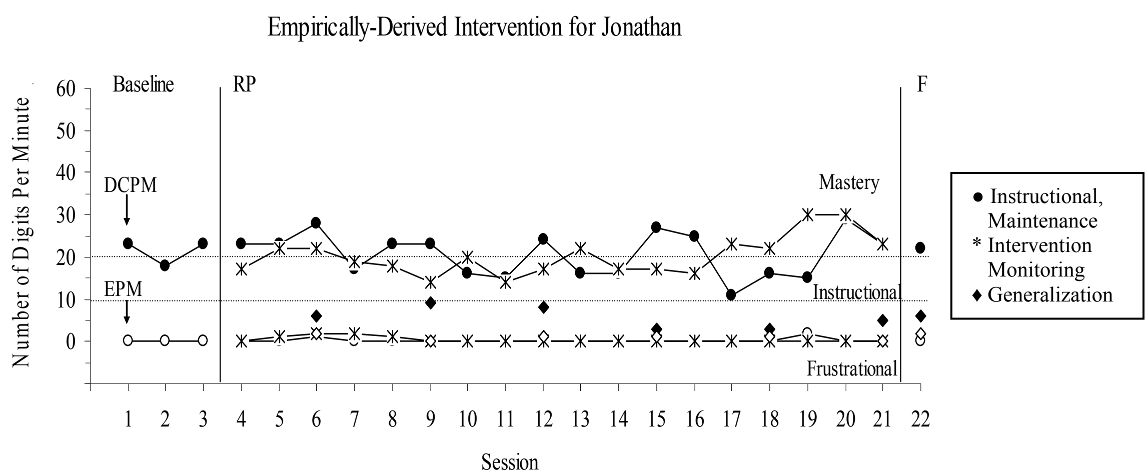

Figure 3.7 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Jonathan

Note. Closed and open data points indicate DCPM and EPM, respectively.
RP = repeated practice; F = follow-up.

Jamario. The FI intervention was selected for Jamario based on the experimental analysis results. Figure 3.8 displays Jamario's calculation performance on instructional worksheets during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the FI intervention. During baseline, Jamario obtained a mean of 20.4 DCPM and 0.8 EPM. He reached mastery level with an increasing trend and high variability during baseline. After the FI intervention was introduced, on intervention monitoring worksheets, Jamario showed an increase in his fluency (*M* = 22.9 DCPM, range = 16-29) and maintained low error rates (*M* = 0.6 EPM, range = 0-2). On maintenance worksheets, he also showed an increase in his fluency (*M* = 21.8 DCPM, range = 17-28) and maintained low error rates (*M* = 0.6 EPM, range = 0-2). Jamario reached mastery level on both types of worksheets. An increasing trend with high variability in DCPM was observed on both types of worksheets. On generalization worksheets, he performed within frustrational level on the average (*M* = 9.2 DCPM, range = 7-11) with low error rates (*M* = 0.3 EPM, range = 0-2). However, he reached instructional level in three sessions. Low level, decreasing trend, and little variability in DCPM were observed on generalization worksheets. EPM was stable with little variability on all types of worksheets. During the follow-up session, Jamario enhanced his fluency (*Mdn* = 32 DCPM) and maintained low error rates (*Mdn* = 0 EPM) on instructional worksheets. He scored within instructional level (15 DCPM) with low error rates (0 EPM) on a generalization worksheet. Jamario's individually-derived estimated weekly slope was 0.6 DCPM (13 weeks from CBA through the intervention phase). As such, he met the criterion of the weekly slope for third graders suggested by Fuchs et al. (1993) and, thus, he was considered to be a responder. Across the experimental and intervention phase (a total of 23 sessions), 16 facts classified as known (i.e., facts responded correctly within 3 seconds) were elevated to mastery facts (i.e., facts responded correctly on an intervention monitoring worksheet for three consecutive sessions), indicating an acquisition rate of 0.70 facts per session; 4 facts classified as unknown (i.e., facts that were responded incorrectly within 3 seconds or those that were not responded within 3 seconds) also became known facts, indicating an acquisition rate of 0.17 facts per session. Thus, the FI intervention increased a ratio of known facts by 4% from CBA (69%) through the intervention phase (73%). Overall, the FI intervention was shown to be effective in enhancing Jamario's instructional skill. He maintained his skill during and after the intervention phase. The FI intervention was found to be minimally effective in enhancing his grade-level skills. Jamario maintained low error rates on instructional-level and grade-level worksheets under the FI intervention.

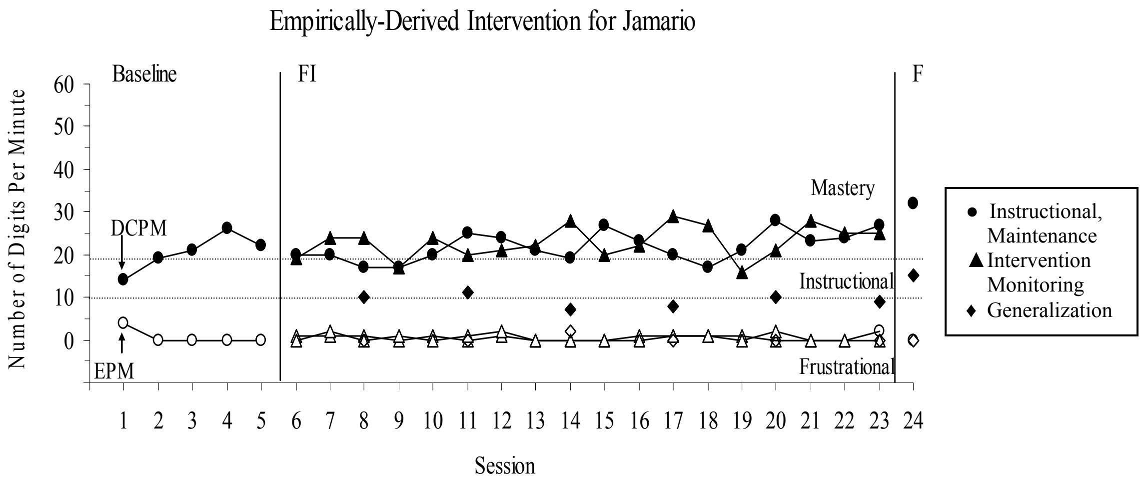

Figure 3.8 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Jamario

Note. Closed and open data points indicate DCPM and EPM, respectively.
FI = folding-in technique; F = follow-up.

Jacklyn. The ICF intervention was selected for Jacklyn based on the experimental analysis results. Figure 3.9 displays Jacklyn's calculation performance on instructional worksheet during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the ICF intervention condition. During baseline, Jacklyn obtained a mean of 20.2 DCPM (range = 13.0 -25.0) and 0 EPM. She reached mastery level with an increasing trend and high variability during baseline. After the ICF intervention was introduced, Jacklyn increased her fluency (*M* = 22.9 DCPM, range = 13-28) and maintained low error rates (*M* = 0.4 EPM, range = 0-2) on intervention monitoring worksheets. She also showed an increase in her fluency (*M* = 23.8 DCPM, range = 18-30) and maintained low error rates (*M* = 0.4 EPM, range = 0-2) on maintenance worksheets. On intervention monitoring and maintenance worksheets, Jacklyn reached mastery level on both types of worksheets. An increasing trend was observed in DCPM on both types of worksheets. Variability in DCPM was less on intervention monitoring than maintenance worksheets (*SD* = 3.1 DCPM and 3.8 DCPM, respectively). On generalization worksheets, she performed within frustrational level (*M* = 8.2 DCPM, range = 6-9) with low error rates (*M* = 1 EPM, range = 0-2). A nearly flat trend with little variability in DCPM was observed on generalization worksheets. EPM was stable with little variability on all types of worksheets. During the follow-up session, Jacklyn enhanced her fluency (*Mdn* = 28 DCPM) and maintained low error rates (*Mdn* = 0 EPM) on instructional worksheets. She scored within frustrational level (8 DCPM) with low error rates (0 EPM) on a generalization worksheet. Jacklyn's individually-derived estimated weekly slope was 0.8 DCPM (13 weeks from CBA through the intervention phase). As such, she met the criterion of the weekly slope for third graders suggested by Fuchs et al. (1993) and, thus, she was considered to be a responder. Overall, the ICF intervention was effective in enhancing Jacklyn's calculation skill. She maintained her skill during and after the intervention phase. The ICF intervention was not effective in increasing her grade-level skills. Jacklyn maintained low error rates on instructional-level and grade-level worksheets under the ICF intervention. It should be noted that an intervention session was conducted with Jacklyn across 4 days once due to her unforeseen events (e.g., family crisis).

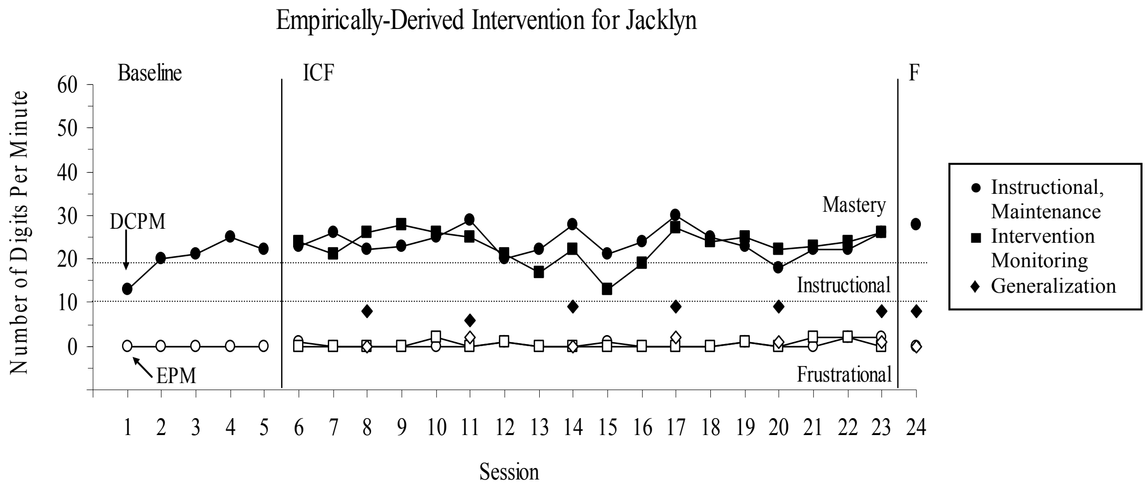

Figure 3.9 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Jacklyn

Note. Closed and open data points indicate DCPM and EPM, respectively. ICF = immediate corrective feedback; F = follow-up.

Intervention Responsiveness of the Students in Special Education

Responsiveness to empirically-derived intervention for student in special education is outlined in this section. Baseline, maintenance, generalization, and follow-up results are also outlined. Decision of responders or nonresponders is outlined. Figure 3.10 to 3.12 display data on instructional, intervention monitoring, and generalization worksheets for students in special education.

Kathy. The RP intervention was selected for Kathy based on the experimental analysis results. Figure 3.10 displays Kathy’s calculation performance on instructional worksheets during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the RP intervention. During baseline, Kathy obtained a mean of 24 DCPM and 0 EPM. She performed at instructional level with an increasing trend and little variability during baseline. After the RP intervention was introduced, on intervention monitoring worksheets, Kathy showed an increase in her fluency (*M* = 29.7 DCPM, range = 24-35.4) and maintained low error rates (*M* = 0.6 EPM, range = 0-5). On maintenance worksheets, she also showed an increase in her fluency (*M* = 28.8 DCPM, range = 22-33) and maintained low error rates (*M* = 0.2 EPM, range = 0-3). She continued to perform at instructional level on both types of worksheets. An increasing trend with little variability in DCPM was observed on intervention monitoring and maintenance worksheets. On generalization worksheets, Kathy frequently performed at frustrational level (*M* = 15.7 DCPM, range = 11-20) with low errors (*M* = 0.8 EPM, range = 0-4). She once reached instructional level on generalization worksheets. An increasing trend with little variability in DCPM was observed on generalization worksheets. EPM was stable with little variability on all types of worksheets. During the follow-up session, Kathy maintained her fluency (*Mdn* 28 DCPM) with low error rates (*Mdn* = 2 EPM) on instructional worksheets. She scored within frustrational level (13 DCPM) with low error rates (2 EPM) on a generalization worksheet. Kathy's individually-derived estimated weekly slope was 1.2 DCPM (9 weeks from CBA through the intervention phase). As such, she met the criterion of the weekly slope for fourth graders (1.20 DCPM) suggested by Fuchs et al. (1993) and, thus, she was considered to be a responder. Overall, the RP intervention was effective in enhancing Kathy's instructional skill and maintaining the skill during and after the intervention phase. The RP intervention was also effective in enhancing her grade-level skills. The RP intervention maintained low error rates on instructional-level and grade-level worksheets. It should be noted that an intervention session was conducted with Kathy across 3 days once and across 4 days once due to her unforeseen events (e.g., absence for doctor appointment).

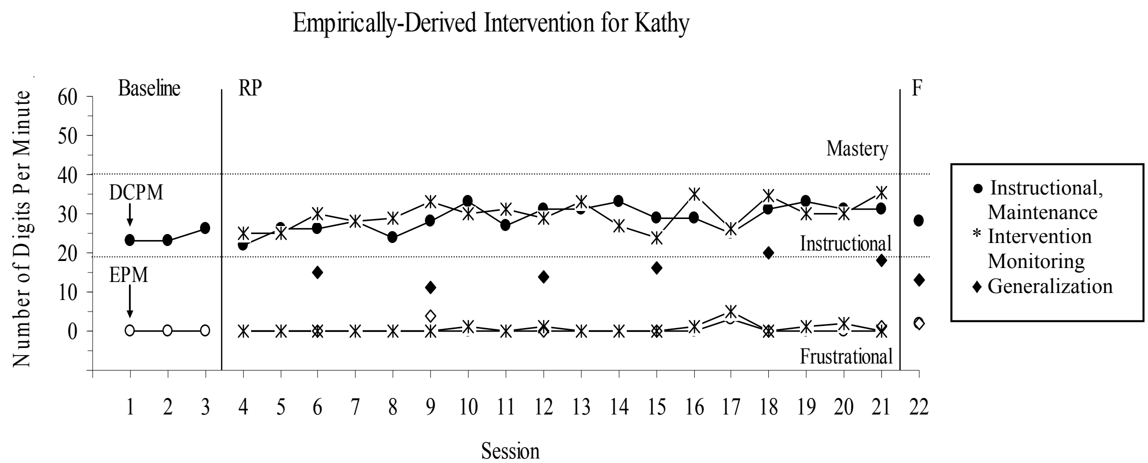

Figure 3.10 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Kathy

Note. Closed and open data points indicate DCPM and EPM, respectively. RP = repeated practice; F = follow-up.

Kaley. The FI intervention was selected for Kaley based on the experimental analysis results. Figure 3.11 displays Kaley's calculation performance on instructional worksheets during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the FI intervention condition. During baseline, Kaley obtained a mean of 23.2 DCPM and 1.2 EPM. She reached instructional level with an increasing trend and high variability during baseline. After the FI intervention was introduced, on intervention monitoring worksheets, Kaley showed a decrease in her fluency (*M* = 19.8 DCPM, range = 15-26), but maintained low error rates (*M* = 1.4 EPM, range = 0-8). On maintenance worksheets, Kaley also showed a decrease in her fluency (*M* = 20.3 DCPM, range = 12-29), but maintained low error rates (*M* = 0.8 EPM, range = 0-2). She continued to perform in a range fluctuating between frustrational and instructional levels on both types of worksheets. An increasing trend in DCPM was observed on both types of worksheets. Variability in DCPM was less on intervention monitoring than maintenance worksheets (*SD* = 2.9 and 4.4, respectively). On generalization worksheets, she performed at frustrational level (*M* = 5.3 DCPM, range = 3-9) with relatively high error rates (*M* = 6.8 EPM, range = 4-11). An increasing trend with little variability was observed on generalization worksheets. EPM was stable with little variability on all types of worksheets. During the follow-up session, Kaley maintained her fluency (*Mdn* = 22 DCPM) and low error rates (*Mdn* = 0 EPM) on instructional worksheets. She scored within frustrational level (7 DCPM) with low error rates (4 EPM) on a generalization worksheet. Kaley's individually-derived estimated weekly slope was 0.5 DCPM (8 weeks from CBA through the intervention phase). As such, she did not meet the criterion of the weekly slope for fourth grade suggested by Fuchs et al. (1993) and, thus, she was considered to be a nonresponder. Across the BEA and intervention phases (a total of 20 sessions), 19 known facts and 1 unknown fact was elevated to mastery facts (a rate of 1 fact per session) and 5 unknown facts became known facts (a rate of 0.25 facts per session). Thus, the FI intervention increased a ratio of known facts by 9.5% from CBA (76.2%) through the intervention phase (85.7%).

Overall, the FI intervention was shown to be effective in enhancing Kaley's instructional skill and maintain her skill during and after the intervention phase. The FI intervention was not shown to be effective in enhancing her grade-level skills. The FI intervention maintained low error rates on instructional-level worksheets, but not on grade-level worksheets.

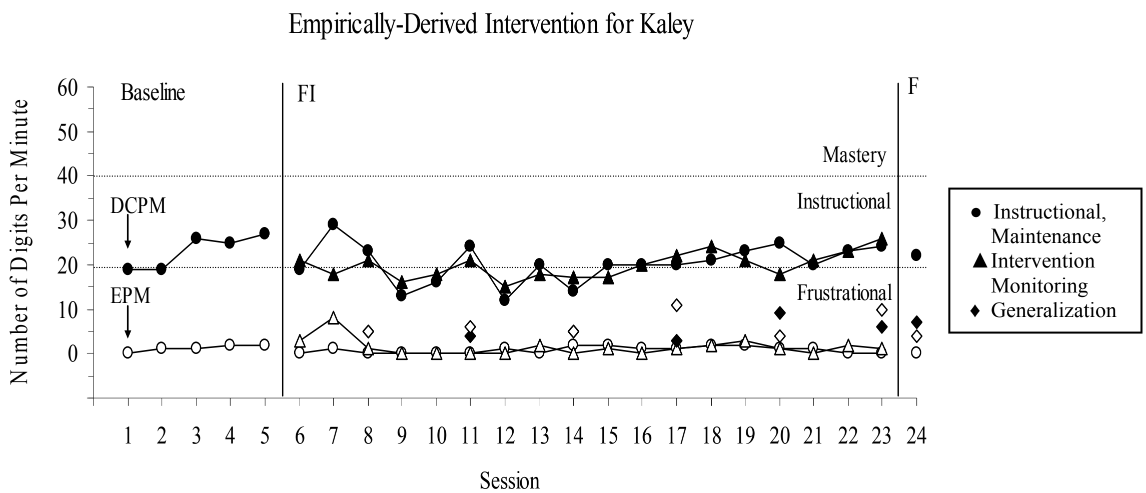

Figure 3.11 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Kaley

Note. Open and closed data points indicate DCPM and EPM, respectively.
FI = folding-in technique.

Tyson. The ICF intervention was selected for Tyson based on the experimental analysis results. Figure 3.12 displays Tyson's calculation performance on instructional worksheets during baseline and while completing intervention monitoring, maintenance, and generalization worksheets under the ICF intervention. During baseline, Tyson obtained a mean of 41 DCPM and 1.9 EPM. He reached mastery level with a decreasing trend and little variability. After the ICF intervention was introduced, Tyson showed a slight increase in his fluency (*M* = 41.8 DCPM, range = 33-46.9) and maintained low error rates (*M* = 0.5 EPM, range = 0-2) on intervention monitoring worksheets. He maintained his fluency (*M* = 41 DCPM, range = 32-51.6) and low error rates (*M* = 0.6 EPM, range = 0-3.5) on maintenance worksheets. Tyson also reached mastery level on both types of worksheets. For DCPM, an increasing trend was observed on intervention monitoring worksheets, whereas a slightly decreasing trend was observed on maintenance worksheets. Less variability in DCPM was observed on intervention monitoring than maintenance worksheets (*SD* = 4.1 and 4.6, respectively). On generalization worksheets, he performed within frustrational level (*M* = 15.3 DCPM, range = 14-18) with low error rates (*M* = 1 EPM, range = 0-3). A nearly flat trend with little variability in DCPM was observed. EPM was stable with little variability on all types of worksheets. During the follow-up session, Tyson maintained his fluency (*Mdn* = 41 DCPM) with low error rates (*Mdn* = 1 EPM). He scored within frustrational level (14 DCPM) with low error rates (2 EPM) on a generalization worksheet. Tyson's individually-derived estimated weekly slope was 0.8 DCPM (12 weeks from CBA through the intervention phase). As such, he did not meet the criterion of the weekly slope for fourth graders suggested by Fuchs et al. (1993) and, thus, he was considered to be a nonresponder. Overall, the ICF intervention was shown to be effective in enhancing Tyson's calculation skill on instructional-level worksheets and maintaining his skill during and after the intervention. The ICF intervention was not shown to be effective in enhancing his grade-level skills. The ICF intervention maintained low error rates on instructional-level and grade-level worksheets.

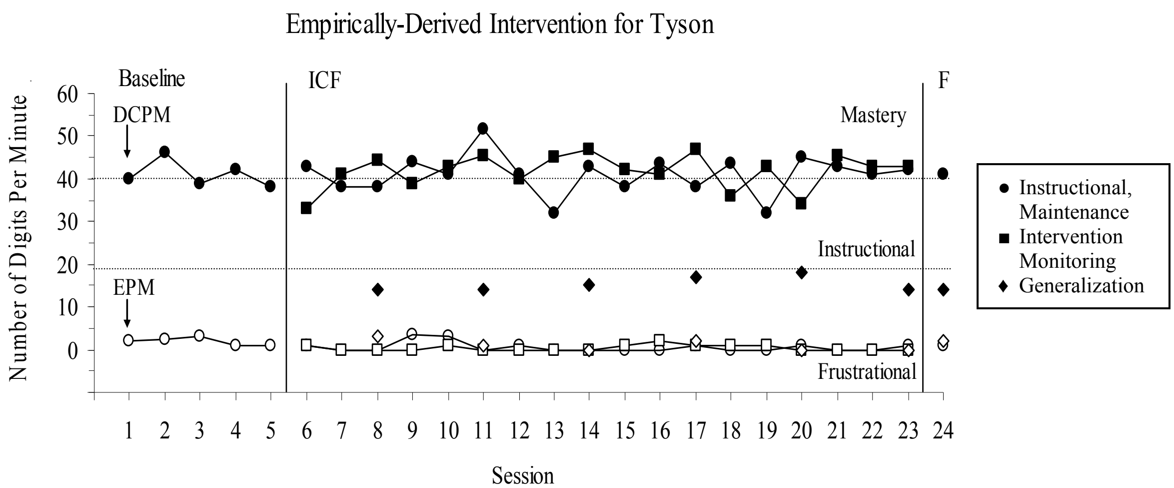

Figure 3.12 Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM) on Instructional, Intervention Monitoring, Maintenance, and Generalization Worksheets during Baseline, Intervention, and Follow-Up Phases for Tyson

Note. Closed and open data points indicate DCPM and EPM, respectively. ICF = immediate corrective feedback; F = follow-up.

Social Validity and Self-Efficacy

Hypothesis Two

Students, with and without SLD in mathematics calculation, will demonstrate satisfaction and self-efficacy following implementation of the intervention that best addressed their skill deficit.

In this section, data regarding Hypothesis Two is presented. Social validity (satisfaction) and self-efficacy was assessed using the *Student Social Validity Checklist* that consisted of the Satisfaction and Self-Efficacy Scales. On each scale, a total score was calculated, indicating the degree of the targeted perception (satisfaction, self-efficacy). Possible range for a total score on each scale was 5 to 15.

The Satisfaction Scale was administered at the end of the intervention phase (i.e., the last intervention session). For satisfaction, students without and with SLD indicated a mean rating of 2.9 and 2.9, respectively (see Table 3.6 and Figure 3.13). Additionally, the mean ratings on the Satisfaction Scale for the RR, FI, and ICF interventions were 3.0, 2.9, and 2.8, respectively. Specifically, all students rated "*True*" on items including "This intervention to improve my math skills was fair," "This intervention is good one to use with other children," and "I think this intervention helps me do better in school." No items were rated as "*Not True*."

Table 3.6 Mean Ratings on the Satisfaction and Self-Efficacy Scales for a Group of Students with and without SLD

Group	Time Measured: Before CBA	Time Measured: Pre-intervention	Time Measured: Post-intervention
	Satisfaction Scale		
General	–	–	2.9
SPED	–	–	2.9
	Self-Efficacy Scale		
General	2.1	2.3	2.7
SPED	2.2	2.3	2.6

Note. The general and special education groups included students without and with Specific Learning Disability (SLD), respectively. CBA = curriculum-based assessment; SPED = special education.

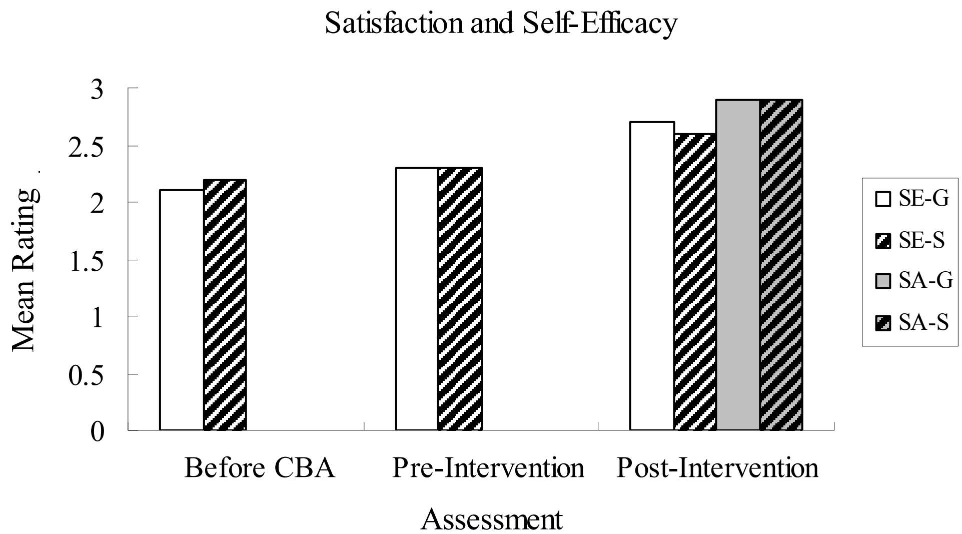

Figure 3.13 Mean Ratings on the *Student Social Validity Checklist, Satisfaction and Self-Efficacy Scales* for a Group of Students in General and Special Education

Note. The general and special education groups included students without and with Specific Learning Disability (SLD), respectively. Bar charts with diagonal lines indicate data for a group of students with SLD. Charts with and without background indicate the data of the Satisfaction and Self-Efficacy Scale, respectively. SA = satisfaction; SE = self-efficacy; G = general education; S = special education.

Table 3.7 Rating Changes on the Self-Efficacy Scale for a Group of Students with and without SLD

Phase	Self-Efficacy Change[a]	Mean Week[b]	Rate[c]
	General Education		
CBA – EA	0.2	6.7	0.03
Intervention	0.4	6.0	0.07
	Special Education		
CBA – EA	0.1	3.7	0.03
Intervention	0.3	6.0	0.05

Note. The general and special education groups included students without and with Specific Learning Disability (SLD), respectively.
CBA = curriculum-based assessment;
EA = experimental analysis; [a] = self-efficacy change;
[b] = mean weeks spent for the phase;
[c] = rate of self-efficacy change.

The Self-Efficacy Scale was administered the following three times across the study: (a) before CBA, (b) at the beginning of the first intervention session, and (c) at the end of the last intervention session. For self-efficacy, ratings by students without SLD showed means of 2.1, 2.3, and 2.7 across the study (i.e., before CBA, first intervention session, and last intervention session, respectively); ratings by students with SLD showed means of 2.2, 2.3, and 2.6 across the study. For both of the groups, self-efficacy increase rate was larger during the intervention phase than during CBA and BEA (see Table 3.7).

Specifically, for the Self-Efficacy Scale, students without SLD showed a weekly increase rate of 0.03 and 0.07 points during CBA and BEA and during the intervention phase, respectively, whereas students with SLD showed a weekly increase rate of 0.03 and 0.05 points during CBA and BEA and during the intervention phase, respectively. Specifically, on one item, “I feel I can figure out the answers almost always,” a group of students without SLD showed a decrease in their mean rating by 0.6 from CBA to the intervention phase; but, then showed an increase (1.0) during the intervention phase. On one item, “I feel I can memorize math problems easily,” a group of students with SLD showed a decrease in their mean rating by 1 from CBA to the intervention phase; but, then showed an increase (0.7) during the intervention phase. Overall, a group of students, with and without SLD, almost equally enhanced their self-efficacy across the study.

CHAPTER IV

DISCUSSION

The purpose of this study was to examine and compare the intervention responsiveness of elementary-aged students, with and without Specific Learning Disability (SLD), for mathematics calculation. In this section, the two hypotheses tested in the study are discussed based on the results. Implications and future research questions are also stated.

Hypotheses

In this section, the two hypotheses tested in this study are discussed based on the results.

Hypothesis One

Hypothesis One stated that students, with and without SLD in mathematics calculation, would demonstrate responsiveness when an effective intervention was identified through a brief experimental analysis (BEA). Specifically, intervention responsiveness was compared for a pair of students with and without SLD who responded to one of the following interventions: (a) repeated practice (RP), (b) folding-in technique (FI), and (c) immediate corrective feedback (ICF).

In this study, for each student, an individually-derived estimated weekly slope of progress from curriculum-based assessment (CBA) through the intervention phase was calculated and examined as to whether the student was a responder based on the aspirational goal of a weekly slope based on the third or fourth grade expectations for growth (Fuchs et al., 1993). Comparison in intervention responsiveness for a pair of students who responded to each intervention (the RP, FI, and ICF interventions) is summarized below.

Comparison of responsiveness to the RP intervention between Jonathan and Kathy. Comparing the intervention responsiveness of Jonathan and Kathy under the RP intervention, both students met the criterion of responders with Jonathan (0.4 DCPM above the critical slope expected for the third grade; Fuchs et al., 1993) exceeding the criterion more than Kathy (at the criterion slope for the fourth grade). Jonathan demonstrated higher levels of performance than Kathy; Jonathan reached mastery level on intervention monitoring and maintenance worksheets, whereas Kathy continued to perform at instructional level on both types of worksheets. On the other hand, Kathy demonstrated less variability than Jonathan (e.g., a standard deviation of 3.5 DCPM and 4.6 DCPM on instructional worksheets for Kathy and Jonathan, respectively). On generalization worksheets, Kathy demonstrated higher levels of performance than Jonathan. Specifically, Kathy reached instructional level, whereas Jonathan continued to perform at frustrational level. Both of the students demonstrated low error rates on all types of worksheets. Overall, Jonathan demonstrated higher responsiveness than Kathy in terms of rates of progress and levels of performance.

Comparison of responsiveness to the FI intervention between Jamario and Kaley. Comparing the intervention responsiveness of Jamario and Kaley under the FI intervention, only Jamario was identified as a responder (0.1 DCPM above the critical slope for the third grade). Jamario demonstrated higher levels of performance than Kaley. For example, on intervention monitoring and maintenance worksheets, Jamario reached mastery level, whereas Kaley performed in a range between frustrational and instructional levels. On generalization worksheets, Jamario reached instructional level, whereas Kaley performed in a frustrational level. Both of the students demonstrated low error rates on intervention monitoring and maintenance worksheets. However, Kaley demonstrated higher acquisition rates of facts than did Jamario (e.g., an acquisition rate of 9.5% and 4% for known facts from CBA through the intervention phase for Kaley and Jamario, respectively). Overall, Jamario demonstrated higher responsiveness than Kaley in terms of rates of progress and levels of performance.

Comparison of responsiveness to the ICF intervention between Jacklyn and Tyson. Comparing the intervention responsiveness of Jacklyn and Tyson under the ICF intervention, only Jacklyn was identified as a responder (0.3 above the critical slope for the third grade). Both of the students demonstrated essentially equal responsiveness in terms of rates of progress and levels of performance. For example, both of them demonstrated an equal individually-derived estimated weekly slope of 0.8 DCPM. They reached mastery level on intervention monitoring and maintenance worksheets. However, both of them performed in frustrational level on generalization worksheets. They maintained low error rates on all types of worksheets. Overall, Jacklyn and Tyson demonstrated equal responsiveness, with only Jacklyn being identified as a responder based on the change in estimated weekly slopes.

Comparison of intervention responsiveness between students with and without SLD. Comparing intervention responsiveness between students with and without SLD as a group, on the average, the third grade students without SLD (*M* = 0.8 DCPM across an average of 12.7 weeks) and the fourth grade students with SLD (*M* = 0.8 DCPM across an average of 9.7 weeks) demonstrated an equal average estimated weekly slope. Based on the criterion of weekly slopes (Fuchs et al., 1993), all of the students without SLD and one student with SLD (i.e., Kathy) were identified as responders. During the intervention phase, on instructional-level worksheets (i.e., intervention monitoring and maintenance worksheets), the three students without SLD reached mastery level, whereas only one student with SLD (i.e., Tyson) reached mastery level. All students maintained their instructional skill on maintenance worksheets. On generalization worksheets, on the average, all students with and without SLD performed at frustrational level, indicating difficulty in generalization from instructional-level to grade-level worksheets in these students. All of the students maintained low error rates on all types of worksheets, except for Kaley, who demonstrated relatively high error rates on generalization worksheets during some sessions. All of the students maintained fluency and low error rates during the follow-up session. Some students without SLD (e.g., Jamario and Jacklyn) showed further increase in fluency during the follow-up session.

In summary, all of the students without SLD, but only one student with SLD (i.e., Kathy) met the criterion of responders. However, when group average estimated weekly slopes were calculated, the students, with and without SLD, demonstrated an equal average estimated weekly slope (.80). This might indicate that there was no essential difference in rates of progress between the two groups. Overall, Hypothesis One was partially supported based on the data for the students' intervention responsiveness.

Hypothesis Two

Hypothesis Two stated that students, with and without SLD in mathematics calculation, would demonstrate higher ratings of satisfaction and self-efficacy following implementation of the intervention that best addressed their skill deficit. Regarding satisfaction, the students with and without SLD demonstrated equally high mean ratings on the *Student Social Validity Checklist, Satisfaction Scale*, regardless of types of interventions. The results indicate that empirically-derived interventions were equally acceptable to the students with and without SLD. Regarding self-efficacy, students with and without SLD increased their mean ratings on the *Student Social Validity Checklist, Self-Efficacy Scale* across the study. Additionally, for self-efficacy, students with and without SLD demonstrated higher mean rating increases during the intervention phase than during CBA and BEA. Specifically, for the Self-Efficacy Scale, students without SLD showed an average weekly increase rate of 0.03 and 0.07 points during CBA and BEA and during the intervention phase, respectively; students with SLD showed an average weekly increase rate of 0.03 and 0.05 points during CBA and BEA and during the intervention phase, respectively. The results indicate that the students' self-efficacy increase was due to the effects of an empirically-derived intervention rather than those of mere completion of worksheets (i.e., practice effects) during CBA and BEA. Overall, Hypothesis Two was supported based on the data for the students' satisfaction and self-efficacy.

Interpretation

Interpretation of the results of the study is summarized in this section.

Idiosyncratic Intervention Responsiveness

First, idiosyncratic responsiveness was observed in the students with and without SLD in this study. Specifically, during the experimental analysis, all students with and without SLD demonstrated DCPM higher under one intervention condition than that under the other two intervention conditions, indicating that intervention responsiveness is idiosyncratic (e.g., Jonathan and Kathy for the RP intervention). Several potential variables may have affected their responsiveness. One of the variables may be specific to a student's classification relative to the Instructional Hierarchy (Daly et al., 1996; Haring & Eaton, 1978) as measured by a ratio of known facts assessed at the beginning of the study. For example, Jamario and Kaley, who responded to the FI intervention, demonstrated a relatively low ratio of known facts (69% and 76.2% for Jamario and Kaley, respectively). This indicates that the two students lacked automaticity in calculation; thus, they may have been in the acquisition stage of the Instructional Hierarchy (i.e., the first stage of the Instructional Hierarchy). On the other hand, the other four students demonstrated a high ratio of known facts. Specifically, both Jonathan and Kathy, who responded to the RP intervention, demonstrated 82.8% of known facts. Jacklyn and Tyson, who responded to the ICF intervention, demonstrated 79.7% and 80%

of known facts, respectively. This indicates that these students had a higher level of automaticity (approximately 80%), but had not mastered some of the facts. Thus, they may have been in transition between the acquisition and fluency (the second stage of the Instructional Hierarchy) stages or in the fluency stage of the Instructional Hierarchy. For students who were in the acquisition stage (e.g., Jamario and Kaley), interventions including drill and practice components such as a flashcard drill (e.g., the FI intervention) may have promoted memorization of mathematics facts. In fact, in the Cates (2005) study, elementary-aged students with calculation deficits demonstrated acquisition of mathematics facts under either flashcard or computer drill. Cates stated one possible reason for these results was that matching the pace of flashcard/computer presentation (e.g., slow/fast pace of flashcard/computer presentation) and a student's rate of fact acquisition (fast/slow acquisition rates of younger/older students) was likely to enhance fact acquisition. Thus, a drill and practice component corresponding to the acquisition rate may effectively enhance calculation automaticity. In the current study, one of the differences between the FI intervention and the other two interventions was that during the FI intervention, students practiced only a limited number of facts (10 facts), many of which (70%) were known facts, whereas during the other two interventions (the RP and ICF interventions), students practiced all facts on a worksheet including a random ratio of known facts. Practice of a limited number of facts including a high ratio of known facts may be an effective component of the FI intervention which promotes students' acquisition of facts. Additionally, the interaction between a student and an interventionist during the FI intervention (e.g., an interventionist's presenting a flashcard and saying, "What is 3+4?" and providing feedback regarding accuracy such as "Correct" or "No, three plus four is seven") may also enhance the student's performance, motivation, and/or attention to tasks. This possible explanation is also supported by the fact that the second most effective intervention identified during the experimental analysis for Jamario and Kaley was the ICF intervention, an intervention that also includes practice and feedback components. Conversely, the RP intervention is an intervention that includes a practice component only with no overt interaction with an interventionist.

On the other hand, for the students who had acquired a high ratio of facts (i.e., Jacklyn, Jonathan, Kathy, and Tyson), interventions including a practice component, with and without a feedback component (e.g., the ICF and RP interventions), may have promoted their accurate performance. It is possible that practice of all facts on a worksheet, rather than a limited number of facts as in the FI intervention, may have promoted their acquisition of all of the facts on the worksheet. In fact, all of these four students demonstrated the lowest level of performance (i.e., mean number of digits correct per minute (DCPM)) under the FI intervention during the experimental analysis. These results suggest that initial ratios of known facts might not have been a better predictor of the responsiveness of students who responded to the RP or ICF intervention. Rather, for these students, some other variables may have influenced their responsiveness. For example, previous research demonstrated that initial levels of performance based on CBM predicted students' responsiveness. Rhymer et al. (1998) demonstrated that in explicit timing (see Van Houten & Thompson, 1976), students with high levels of baseline performance demonstrated higher levels of accuracy than did those with middle and low levels of baseline performance. Codding et al. (2007) extended Rhymer et al. by demonstrating that students whose initial level was within frustrational level (i.e., those who lacked skills) responded better to Cover, Copy, and Compare (see Skinner et al., 1989), an intervention including a practice component. Conversely, those whose initial level was within instructional level (i.e., those who had a certain level of skills) responded better to explicit timing, an intervention including both practice and feedback components. Codding et al. explained that a practice component included in both interventions may have enhanced the calculation skills of students in both groups. However, a feedback component in explicit timing may also have enhanced the motivation of students whose initial level was within instructional level rather than students whose initial level was within frustrational level.

Similar findings were also observed in the present study. For example, Jonathan and Kathy, who responded to the RP intervention (i.e., an intervention including a repeated practice component), scored within a range fluctuating between frustrational and instructional levels during the CBA (10 and 19 DCPM for Jonathan and Kathy, respectively). This finding may indicate that the two students had acquired skills, but had difficulty demonstrating automaticity and, thus, the RP intervention yielded their highest responsiveness. On the other hand, Jacklyn and Tyson, who responded to the ICF intervention (i.e., an intervention including both practice and feedback components), scored within instructional level during the CBA. This finding may indicate that these two students had a minimum level of automaticity and, thus, the ICF intervention yielded their highest responsiveness. It should be noted that during the experimental analysis, for

Jonathan and Kathy, the ICF intervention was identified as the second most effective intervention, whereas for Jacklyn and Tyson, the RP intervention was identified as the second most effective intervention. Additionally, for all of the four students, the FI intervention was identified as the least effective intervention (see Table 3.1 for experimental analysis data). These results may also indicate that the four students had acquired a higher level of facts (e.g., approximately 80% of known facts) and, thus, practice of all facts in a worksheet, rather than a limited number of facts as in the FI intervention, was an effective component to enhance their automaticity. Notably, Jamario and Kaley, who responded to the FI intervention, scored within instructional and frustrational level during the CBA, respectively. A potential indication is that for these students, initial levels of performance based on CBM might not have a better predictor of their responsiveness. It should also be noted that the three interventions are different in terms of the student's independence (e.g., an interventionist's commitment, interaction between the student and interventionist, etc.). For example, a student with a high initial level of performance may respond to an intervention that requires less or no interventionist's commitment (e.g., feedback) and/or interaction with an interventionist (e.g., flashcard drill) such as the ICF and RP interventions.

Intervention Responsiveness and Maintenance of Skills

Second, this study examined the effects of empirically-derived interventions on students' instructional skills. Maintenance of intervention effects was also assessed. Follow-up assessment was conducted one week after the termination of an intervention. All students reached instructional or mastery level on intervention monitoring and maintenance worksheets when they received an intervention. Each of the students demonstrated almost equal levels of fluency and low error rates on both intervention monitoring and maintenance worksheets, indicating that the skill acquired through an intervention was well maintained. Some students with and without SLD (i.e., Jamario, Jacklyn, Kaley, Kathy) also demonstrated an increasing trend on intervention monitoring worksheets. Furthermore, some students with and without SLD (i.e., Jonathan, Kaley, Tyson) demonstrated less variability (as measured by standard deviation) on intervention monitoring than maintenance worksheets. This indicates that an empirically-derived intervention was effective in stabilizing their calculation performance within a certain level (e.g., instructional or mastery).

However, large increases in level were not observed from baseline to the intervention phase for all students. This might have been due to ceiling effects. That is, because the students enhanced their level when they received interventions during the experimental analysis, it might have been difficult for them to further increase their level on intervention monitoring and maintenance worksheets during the intervention phase. Another possible reason is that students might have been more cautious and sought to perform accurately on intervention monitoring and/or maintenance worksheets after they received interventions. For example, from the experimental analysis through the intervention phase, under the ICF intervention, Tyson maintained a stable level of performance, but demonstrated lower error rates during the two phases than during baseline. Fatigue after practicing several worksheets during an intervention might be also a possible factor that affected students' performance on intervention monitoring and maintenance worksheets (e.g., Jonathan for the RP intervention, who slightly decreased a level of performance on intervention monitoring and maintenance worksheets after the RP intervention was introduced while maintaining low error rates). All students with and without SLD also maintained or further increased their level during the follow-up session. Overall, empirically-derived interventions were effective in enhancing and maintaining the instructional skills of the students.

Generalization of Skills

Third, generalization from instructional-level to grade-level skills was also examined in this study. Most of the students with and without SLD did not generalize their skill on grade-level worksheets. Only Jamario and Kathy reached instructional level on generalization worksheets. For Jamario (third grade), it may be because his instructional level was at the second grade rather than at the first grade and, thus, generalization in skills was likely to occur across the close grades (i.e., second and third grades), compared to the other students whose instructional level was at the first grade. However, on generalization worksheets, Jamario performed within frustrational level on the average. Jacklyn and Tyson also reached a range between frustrational and instructional levels during some sessions. On the other hand, Jonathan and Kaley performed within lower frustrational level during some session. In terms of the Instructional Hierarchy, for all students, their levels were still within the acquisition and/or fluency stages, rather than the generalization or adaptation stage, after they received an empirically-derived intervention.

Another possible reason for students' difficulty in obtaining generalization is due to differences in types of skills between intervention monitoring and generalization worksheets (e.g., addition with and without regrouping, addition with different number of digits). Specifically, in this study, an instructional skill included one-digit plus one-digit addition (single-digit addition sums to 10 or 18), whereas generalization skills included two- to three-digit addition with and without regrouping for the third grade and five- to six-digit addition with and without regrouping for the fourth grade. Researchers have stated that generalization is likely to occur when the two stimuli, between which generalization is to occur, are the same or close (e.g., Stokes & Baer, 1977). For mathematics skills, researchers may examine whether differences in types of skills affect students' generalizability (e.g., addition without regrouping to those with and without regrouping, one-digit addition to multiple-digit addition without regrouping). It should be noted that some or all students in this study might have been slow learners who needed more intervention opportunities to develop instruction-level fluency and to generalize these skills to grade-level materials. In fact, some students demonstrated an increasing trend on generalization worksheets (e.g., Kaley and Kathy). Some students also demonstrated progress on grade-level skills, when compared to their initial level of performance assessed during CBA. For example, from CBA through the intervention phase, Kathy and Tyson increased their ability level by 5.7 and 7.3 DCPM on fourth-grade worksheets, respectively. This growth might be due to intervention effects rather than extraneous variables, given the short duration between the two phases (i.e., approximately one month). The increases are also meaningful, given the students' history of academic delays (e.g., retention). However, further interventions may have been necessary to enhance their grade-level skills. Overall, difficulty in generalization of skills from instructional-level to grade-level materials may be one of the characteristics of students with mathematics difficulties.

Satisfaction and Self-Efficacy

Fourth, this study examined the effects of empirically-derived interventions on students' satisfaction and self-efficacy. In this study, students demonstrated high levels of satisfaction with intervention procedures. Although some researchers have found social validity for mathematics interventions which were not empirically-derived (e.g., Arra & Bahr, 2005; DuPaul et al., 1998), the results of the present study indicated that empirically-derived interventions implemented in this study were both effective and acceptable. However, this fact should be further examined in future research.

The present study also demonstrated that empirically-derived interventions equally enhanced the self-efficacy of students with and without SLD. Students with and without SLD demonstrated higher mean rating increases (i.e., two to three times larger increase) in self-efficacy during the intervention phase than found during CBA to the intervention phase (see Table 3.6). The larger amount of intervention opportunities during the intervention phase (18 sessions) than during CBA to the intervention phase (4 sessions for Kaley and 13 to 14 sessions for the other students) might have been responsible for the higher self-efficacy increase. Notably, analysis of the change in the self-efficacy of each student across the study revealed that three students demonstrated larger increases during the intervention phase; one student maintained the same level across the study, and two students demonstrated higher increases during CBM and BEA. The results may indicate that an empirically-derived intervention implemented during the intervention phase enhanced or maintained students' self-efficacy relative to both empirically- and non-empirically-derived interventions that were alternatively implemented during the experimental analysis.

Summary

In summary, intervention responsiveness was idiosyncratic for students with and without SLD in this study. Specifically, during experimental analysis, one intervention produced a higher DCPM than did the other interventions for all students. Additionally, an empirically-derived intervention was effective in enhancing students' instructional skill and maintaining their skill during the intervention phase. However, students without SLD were likely to respond to an empirically-derived intervention compared to those with SLD. In fact, all students without SLD, but only one student with SLD were identified as responders based on the ambitious weekly slope criterion for third or fourth graders suggested by Fuchs et al. (1993). Empirically-derived interventions were not effective in enhancing students' skills on generalization worksheets. However, some students demonstrated slow progress on generalization worksheets after receiving empirically-derived interventions (e.g., Jamario, Kathy, and Tyson). For satisfaction and self-efficacy, students with and without SLD showed equally high mean ratings on the Satisfaction Scale. Students with and without SLD also showed an increase in mean ratings on the Self-Efficacy Scales across the study. These findings indicate that empirically-derived interventions were equally acceptable to students with and without SLD and effective in enhancing self-efficacy in these students.

Implications and Relevant Future Research

This study is noteworthy in that intervention responsiveness was compared between elementary-aged students with and without SLD in the area of mathematics calculation using a single subject design, sensitive to each student's idiosyncratic responsiveness, to determine differential responsiveness characteristics of responders to three empirically-derived interventions (i.e., RP, FI, and ICF). In this section, implications and related future research which will further address these implications are stated.

Intervention Responsiveness

Responsiveness may be analyzed in terms of rates of progress (weekly slopes) and levels of performance. In the study, average estimated weekly slopes were compared between groups of students with and without SLD. On the average, third-grade students without SLD and fourth-grade students with SLD demonstrated an equal mean weekly slope. Based on the aspirational goal of weekly slopes found in previous studies (Fuchs et al., 1993), all students without SLD and one student with SLD (i.e., Kathy) were identified as responders. Regarding levels of performance, during the intervention phase, the three students without SLD and one student with SLD (i.e., Tyson) performed within high instructional and mastery levels on intervention monitoring and maintenance worksheets. On the other hand, one student with SLD (i.e., Kathy) continued to perform at instructional level and another student with SLD (i.e., KB) continued to perform in a range between frustrational and instructional levels. Superior responsiveness of students without SLD to those with SLD was also found in previous studies using a group design in general mathematics (calculation and reasoning) (e.g., Fuchs et al., 2004). However, the present study extended the previous studies by identifying an intervention experimentally, rather than selecting an intervention a priori (e.g., Fuchs et al., 1990, 2004). In the present study, individual responsiveness differences were also examined using a single subject design rather than a group design (e.g., Fuchs et al., 2004). Furthermore, the present study demonstrated that rates of progress (e.g., weekly slopes) and levels of performance are viable measures that quantify responsiveness, which can be used in the examination of students' responsiveness in the Response-to-Intervention (RtI) model. Despite the findings of this study, further research is warranted to examine responsiveness difference in students with and without SLD in mathematics to confirm the findings of this study (e.g., replication of this study to examine whether responsiveness is likely to be higher for students without SLD than those with SLD in the same or close grades and whether differences are found between students with and without SLD in terms of satisfaction and self-efficacy after receiving empirically-derived interventions, etc.).

Measured Variables

A second implication is that in this study, some measured variables may have also predicted students' idiosyncratic intervention responsiveness (e.g., initial ratio of known facts, initial CBM level). Based on the results of the study, the following issue is assumed: For students whose initial ratio of facts is low, initial ratios of facts, rather than initial levels of performance based on CBM, primarily predicts their responsiveness; for these students, acquisition of a limited number of facts including a high ratio of known facts should be addressed in an intervention. On the other hand, for students whose initial ratio of known facts is high (approximately 80%), initial levels of performance based on CBM, rather than initial ratio of known facts, primarily predicts their responsiveness; for these students, enhancement of automaticity on all facts (e.g., approximately 80% of known and 20% of unknown facts) should be addressed in an intervention. The results of the study also indicate that characteristics that primarily predict responsiveness (e.g., initial ratio of known facts) are idiosyncratic for each student, and a student's intervention responsiveness may be maximized when an intervention addresses specific idiosyncracies. Future research may examine potential variables that predict students' responsiveness to specific interventions. To do so, researchers may use appropriate methodologies, depending on their research purpose, such as a single subject design for identifying *idiosyncratic* aspects of responders or a group design (e.g., statistical procedures) for identifying *common* aspects of responders.

Prior to Special Education

A third implication is that in the RtI model, interventions are implemented with students with academic difficulties and their responsiveness is examined before their special education eligibility is determined. This study is considered a preliminary study which examined the utility of the RtI model when implemented in school settings. In the study, for each student, an intervention was selected empirically and the effects of the intervention were examined for a specified period of time (e.g., 6 weeks). Based on the results of this study, empirically-derived interventions within the RtI model may enhance the responsiveness of students who are academically at-risk (e.g., students with and without SLD). In fact, in this study, some students with SLD in mathematics calculation (e.g., Kathy and Tyson) demonstrated a rate of progress (e.g., an individually-derived estimated weekly slope of progress exceeding the aspirational goal for their grade) and/or a level of performance (e.g., mastery level) after they received empirically-derived interventions; and, thus, these students might not have been identified with SLD (e.g., within a dual discrepancy model in which only students with low levels of performance and low rates of progress are identified with SLD; Fuchs & Fuchs, 1998).

The present study also demonstrated that BEA is a viable method to identify an effective intervention that is implemented for the examination of a student's responsiveness within the RtI model. However, further research is warranted to examine the utility of BEA within the RtI model. For example, in the present study, whether an intervention identified as most effective during BEA enhanced a student's skills during a subsequent intervention phase was examined; based on the results, empirically-derived interventions were effective in enhancing calculation skills in all students without SLD and one student with SLD (i.e., Kathy). This study also examined whether BEA and extended analysis identified the same intervention as most effective by implementing some or all of the interventions identified as effective during the BEA (e.g., the first and second most effective interventions) during the extended analysis; based on the results, for all students, the same intervention was identified as most effective during BEA and extended analysis (for Jacklyn, the FI and ICF interventions were identified as most effective during BEA, but the ICF intervention was selected based on the results of the BEA and extended analysis). Overall, selection and implementation of potential interventions using reliable methodology (e.g., BEA) is crucial in the examination of students' responsiveness in the RtI model.

Strategy Use

A fourth implication is that some characteristics of students with mathematics deficits, which were found in previous studies, were also observed in this study. For example, in this study, all students with and without SLD lacked automaticity in calculation and used primitive strategies (e.g., finger and/or verbal counting; Gersten et al., 2005; Pellegrino & Goldman, 1987). Some of the fourth-grade students with SLD performed lower than third-grade students without SLD (Russell & Ginsburg, 1984). For example, for the students paired to receive the FI intervention (Jamario in the third grade and Kaley in the fourth grade), Kaley's instructional level (first grade) identified during CBA was lower than that of Jamario (second grade). Kaley, with an individually-derived estimated weekly slope of 0.5 DCPM and frustrational to instructional level performance, also demonstrated a rate of progress and level of performance which were lower than did Jamario, with an individually-derived estimated weekly slope of 0.6 DCPM and mastery level performance. Conversely, some of the characteristics of students with SLD in mathematics found in previous studies were not observed in this study. For example, in previous studies, students with mathematics deficits (e.g., SLD in mathematics) demonstrated frequent errors on simple calculation facts (Russell & Ginsburg). In the present study, however, frequent errors were not observed in the performance of either the students with or without SLD. This may be because the students had basic mathematics skills (e.g., counting skills), which are prerequisite to calculation skills. In previous studies, students with comorbidity of SLD in mathematics and reading used more primitive strategies and made more errors on mathematics calculation problems than did those with SLD in mathematics only (Jordan & Hanich, 2000, 2003; Jordan et al., 2003). Other studies have shown that students with SLD in mathematics also demonstrated lower responsiveness than did those without SLD, even after they received an intervention (e.g., Fuchs et al., 2004). In the present study, however, students with comorbidity of SLD in mathematics and reading (i.e., Kathy and Tyson) demonstrated progress and maintained low error rates when they received an intervention. The students paired to receive the ICF intervention (Jacklyn and Tyson) reached mastery level and demonstrated an equal individually-derived estimated weekly slope.

These results indicate that students' deficits in mathematics calculation determined during a single assessment opportunity may not necessarily predict students' intervention responsiveness. These findings also indicate the necessity of selection of an intervention in an empirical manner and examination of intervention responsiveness, rather than assessment of performance at a single testing opportunity, in the identification of SLD. Empirical selection of an intervention and implementation of the intervention for a period of time (e.g., 6 weeks) may reliably differentiate students who truly need special education services that address their academic deficits and those who can be still taught in general education with individual accommodations (e.g., individually-tailored interventions). Based on the findings of the present and previous studies, lack of automaticity and use of primitive strategies may be common characteristics of students with calculation deficits.

Self-Efficacy

A fifth implication regarding self-efficacy is that in previous studies, students with academic deficits (e.g., those with SLD in mathematics calculation) in middle and high schools demonstrated lower levels of self-efficacy than those without academic deficits (Hampton & Mason, 2003; Lackaye et al., 2006), which may have been due to their unsuccessful learning history. However, in the present study, students with and without SLD demonstrated almost equal initial levels of self-efficacy and almost equal increase rates in self-efficacy across the study. The finding of the present study might be due to the fact that students with SLD in the study had minimum levels of skills (e.g., at least first grade level skill) and, thus, were more confident about their skills than those with significantly low levels of skills (e.g., kindergarten level skills). Additionally, the students with SLD in the study were elementary-aged, rather than those in middle and high schools as in the previous studies and, thus, they were likely to demonstrate an equal level of self-efficacy as those without SLD in the study due to their shorter history of academic delays compared to middle and high school students. Given that self-efficacy predicts students' learning engagement and future achievement (Pietsch et al., 2003; Stevens et al., 2004), early interventions that enhance both academic skills and self-efficacy may be implemented with elementary-aged students.

The intervention responsiveness, satisfaction, and self-efficacy of students who are academically at-risk in mathematics will be further examined in future research (e.g., a longitudinal study).

Limitations and Relevant Future Research

This study has a number of limitations. Threats to internal and external validity are stated below. Future research that may address these limitations is also stated.

Threats to Internal Validity and Relevant Future Research

Extraneous variables. First, some extraneous variables might have affected students' intervention responsiveness in this study (e.g., maturation, on-going mathematics instruction in class, etc.). For example, the students with SLD in this study attended a mathematics class in special education, in which they received an intense mathematics instruction either in a small group or individually. This might have also enhanced their calculation skills at instructional and/or grade level. However, for all students, this effect might have been minimal, given immediate DCPM increase under an empirically-derived intervention during BEA and maintenance of the effects of the empirically-derived intervention during the intervention phases.

The implementation of the AB design in this study was also a threat to internal validity in that the effects of an empirically-derived intervention were not confirmed using subsequent replicated phases. However, for each student, similar intervention responsiveness observed during the experimental analysis and intervention phase may indicate that the effects of the empirically-derived intervention were relatively stable, which in turn may minimize threats to internal validity. In future research, a design including replicated phases may be implemented to confirm intervention effects, when students' intervention responsiveness is examined.

Retention. Second, in this study, all of the students with SLD in the fourth grade had repeated previous grades due to their academic failure in some subject areas (e.g., mathematics, reading, English Language Arts, etc.) and, thus, they had been exposed to academic stimuli more than typical fourth grade students without repetition of previous grades. This exposure to academic stimuli in repeated grades might have affected their initial levels of performance and/or intervention responsiveness. However, the students' immediate level increase from CBA to BEA and maintenance of respective increases in level across all phases indicate that their responsiveness may have been due to intervention effects rather than extraneous variables (e.g., previous education). Future research may compare intervention responsiveness between students with and without SLD at the same grade or at close grades (e.g., third and fourth grades), who did not repeat previous grade.

Experimental analysis. Third, several limitations regarding experimental analysis should be also stated. For example, a period of time spent for an experimental analysis was different for each student; for students for whom BEA was conducted for a longer or shorter period of weeks, a weekly slope might have been underestimated or overestimated, respectively. For example, for a pair of Jonathan and Kathy, an experimental analysis was conducted for 12 and 9 weeks, respectively. However, the two students demonstrated almost equal level increase across the study (i.e., 10.2 and 10.7 DCPM for Jonathan and Kathy, respectively; see Table 3.6). Although Kathy was identified as a responder based on her weekly slope of 1.2 DCPM (=10.7 DCPM increase/9 weeks), her individually-derived estimated weekly slope would have been 0.9 DCPM (=10.7 DCPM/12 weeks) if the experimental analysis had been conducted for 12 weeks, which was an equal period of weeks spent for Jonathan, and she had demonstrated the same estimated weekly slope. This estimated weekly slope of 0.9 DCPM was also below the aspirational goal of 1.2 DCPM for the fourth grade suggested by Fuchs et al. (1993). On the other hand, for students whose initial level was high, individually-derived estimated weekly slopes might be underestimated. For example, for Tyson, because his initial level was within a high instructional level, it might have been difficult for him to demonstrate high increase that may yield a high estimated weekly slope (i.e., ceiling effects). In fact, he reached a mastery level under ICF in the experimental analysis and the intervention phase, and thus may have been a responder in the dual discrepancy model, in which only students who demonstrate both a low level of performance and a low rate of progress are identified as a nonrepsonder. (Fuchs & Fuchs, 1998). This also indicates that the intervention should have been continued to further examine his responsiveness in the RtI model. For example, the intervention may have been implemented using worksheets for the next benchmark and changes in a level of performance and a slope of progress for the benchmark may have been analyzed. This procedure may have been repeated until he is identified as a nonresponder for a certain benchmark which is below his grade level or may have been continued if he had continued to demonstrate progress across benchmarks.

In future research, interventions should be implemented for an equal period of time when responsiveness is compared between students with and without SLD, which may yield more reliable comparison between the two groups. Classification of responders may be also determined based on both levels of performance and slopes of progress (i.e., dual discrepancy analysis).

Different amount of intervention opportunities in the experimental analysis between students for whom an extended analysis was conducted and for whom it was not conducted might have affected their responsiveness in the subsequent intervention phase. For example, although Kaley demonstrated higher responsiveness under FI twice than the other interventions, FI was not effective enough to further enhance her responsiveness in the intervention phase. Kaley also demonstrated a higher mean DCPM during baseline than during the BEA and the intervention phase. She continued to perform within a border range between frustrational and instructional levels during the intervention phase. This might be because she had pervasive mathematics deficits (i.e., SLD in both mathematics calculation and reasoning) as opposed to the other students with SLD only in mathematics calculation and, thus, did not respond well to an empirically-derived intervention (i.e., the FI intervention). However, this might be also because only four sessions in the BEA might not have been sufficient to reliably identify the most effective intervention for Kaley; a different intervention might have been identified as effective if an extended analysis had been conducted. On the other hand, for students who received interventions in an extended analysis, additional opportunities to respond in the extended analysis (e.g., receiving an empirically-derived intervention and the other two interventions, repeatedly completing worksheets under these interventions) might have further enhanced their responsiveness. Additionally, for Jacklyn, because in the BEA, she demonstrated an equal highest DCPM under the two interventions and these two interventions were replicated, a total session of the experimental analysis was one session higher than that of the students for whom an extended analysis was conducted. Previous research demonstrated that repeated practice of academic skills in assessment sessions also enhanced students' performance (e.g., practice effects) (Lee & Tingstrom, 1994; Skinner & Shapiro, 1989). Based on the fact, non-empirically-derived interventions in the experimental analysis might also have enhanced the students' skills.

In future research, an extended analysis may be conducted to all students to confirm the effects of an empirically-derived intervention identified in a BEA. Additionally, as in this study, the intervention that produced the highest mean DCPM across all experimental analysis sessions may be selected and implemented in the subsequent intervention phase.

FI intervention. Fourth, in the FI intervention, 10 flashcards (7 known and 3 unknown facts) were used for fact practice in each intervention session; however, each intervention monitoring worksheet also included 70% of known facts (i.e., 21 problems) and 30% of unknown facts (i.e., 9 problems). As such, difficulty of each worksheet might have been slightly different. Generally, for students most of whose known facts are easy facts (e.g., facts including only lower numbers such as 0, 1, 2, etc.), an intervention monitoring worksheet might be easier than typical worksheets used in other interventions. However, Jamario and Kaley knew 60 to 80% of the facts and, thus, this concern may be ignorable. Additionally, because worksheets always included 70% of known facts and 30% of unknown facts, it might have been difficult for the students to demonstrate rapid progress (e.g., a steep slope). In fact, Jamario and Kaley demonstrated lower slopes than the other students. However, their progress in the FI intervention can be also measured by ratio change of known facts across the study. Specifically, across the study, Jamario and Kaley demonstrated 4% and 9.5% increases in known facts. Jamario and Kaley also demonstrated an acquisition rate of 0.70 and 1.0 facts per session. The results indicate that FI promoted their fact acquisition. Use of worksheets including 70% of known and 30% of unknown facts in FI might be also a limitation in that it is unclear whether students' progress was due to flashcard drill, the ratio of known and unknown facts in the worksheets, or a combination of both. In future research, worksheets including 10 facts used in the intervention plus randomly-selected additional problems may be used in FI. Students' responsiveness may be also compared using worksheets including a constant ratio of known and unknown facts, as used in this study, and typical worksheets including 10 facts that are used for practice in the intervention and additional randomly-selected problems. Furthermore, known and unknown facts may be determined based on students' responses to 10 flashcards used in fact practice rather than their responses on intervention monitoring worksheets administered during the intervention. In this study, of the 10 facts used in the intervention, un-attempted facts on an intervention monitoring worksheet were kept in the same pile of known or unknown facts. However, some of the known or unknown facts might have changed to a different type of facts (e.g., from unknown to known or vice versa) if they had been attempted. This might have also delayed a rate of fact acquisition for the students in the FI intervention (e.g., a rate of change from unknown to known facts across intervention sessions). Even so, in the FI intervention, practice of 10 facts including 7 known and 3 unknown facts may be an effective component to enhance the fact acquisition of students whose initial known ratio is low (e.g., 60 to 80%) as observed in this study.

Number of problems on a worksheet. Fifth, Kathy and Tyson, fourth grade students with SLD, completed an intervention monitoring and maintenance worksheet before 1 minute elapsed for 6.1% and 26.5% of all experimental analysis and intervention sessions, respectively. Although worksheets were similar to those used in typical third and fourth grade classes so that generalization is likely to occur from intervention to classroom materials, 35 to 40 problems might have been adequate task demands for these students. This might be because they repeated previous grades before the study and/or they mastered facts for their instructional level faster than the other students across the study.

In future research, it may be examined whether different task demands (e.g., worksheets including 30 problems versus those including 40 problems) affect students' responsiveness.

Inconsistent time schedule of intervention implementation. Sixth, the inconsistent time schedule of intervention implementation in school settings is also a limitation that might have affected students' performance in interventions. For example, an intervention was implemented with a pair of students with and without SLD for a certain intervention, who stayed at different schools, on different days in some sessions due to unforeseen events including students' absences (e.g., absences due to sickness, doctor appointment, family crisis, etc.) and school events (e.g., tests, school trip, field day, etc.). An intervention was also implemented twice on the same day, with a break between the two implementations (e.g., 30-minute break), with some students in some sessions to make up their previous missed sessions. Given that data collection was conducted at five schools, some students received an intervention often in the morning (e.g., a pair for FI and ICF), whereas the others often in the afternoon (e.g., a pair for RP). Additionally, even the same student received an intervention at different times across the study such as in the morning in some sessions, but in the afternoon in the other sessions, due to the inconsistent schedule of the student and an interventionist (e.g., a school meeting for an interventionist).

Another issue is that an intervention was implemented with some students before or after physical education class in some sessions and this might have affected students' performance and motivation. For example, for Jonathan, an intervention was always implemented immediately before physical education class, which his teacher stated was best time to implement an intervention so that the student would not miss classes in main subject areas (e.g., mathematics, reading, etc.). However, because Jonathan liked physical education class, completion of worksheets in the intervention might have been positive reinforcement for him such that he could go to physical education class soon if he completes worksheets quickly. On the other hand, an intervention was implemented with Jamario and Jacklyn, who stayed in the same class, after physical education class in some sessions. Jamario seemed to be somewhat tired in some sessions, which might have led his lower performance than usual (e.g., the tenth session in an experimental analysis), whereas Jacklyn was consistently motivated and rarely showed tiredness. Inconsistent time schedule of intervention implementation may be a limitation that is specific to natural settings as opposed to contrived settings (e.g., university clinic) in which a researcher or practitioner and a client usually schedule intervention sessions for a consistent time (e.g., a certain time on a certain day each week).

Social validity checklists. Seventh, in this study, to assess students' satisfaction and self-efficacy, social validity checklists were developed from the scales that were used in previous studies. These research-based scales have been demonstrated to be reliable and valid in the assessment of social validity. Although the scales used in the present study were assumed to have a certain level of reliability and validity due to the high psychometric properties of the original scales, the psychometric properties of the scales used in the present study are unknown.

Examination of the psychometric properties of the scales used in this study and improvement of the scales, if satisfactory levels of properties are not obtained, should be conducted in future research.

Interobserver agreement for integrity. Finally, interobserver agreement (IOA) for integrity was assessed only for 6.3% of the sessions across all phases. Low IOA might threats reliability of integrity assessment. However, 100% integrity obtained in all sessions in which IOA was assessed and 100% agreement in the sessions may indicate that interventions were implemented as intended. Immediate level increase from CBA to an experimental analysis and maintenance of the increased level across the subsequent phases observed in all students may also indicate that their progress was due to intervention effects rather than extraneous variables.

Threats to External Validity and Relevant Future Research

Threats to external validity (e.g., limited generalizability of the results of the study across populations, settings, skills such as subtraction and multiplication skills, etc.) are also concerned in this study. For example, only one pair for each intervention was included in this study. Although more students without SLD were identified as responders than those with SLD in the study, this fact should be further examined in future research. Additionally, this study should be replicated with diverse populations. For example, responsiveness may be examined with students with SLD who are identified in other types of discrepancy model (e.g., IQ-achievement discrepancy model). Responsiveness may be examined with students with and without SLD in younger grades (e.g., first and second grades) for the purpose of early prevention of mathematics deficits. Responsiveness differences may be also compared between students who are academically at-risk in general education and those with SLD in older grades (e.g., fifth and sixth grades). Given that the Matthew Effect (Stanovich, 1986) in reading may also occur in mathematics, students with calculation deficits in older grades might demonstrate lower responsiveness on intervention monitoring and generalization worksheets than that observed in the third and fourth grade students in this study. Furthermore, responsiveness may be examined with students with SLD in mathematics reasoning or elementary-aged students with deficits in kindergarten level skills (e.g., number sense, base ten concepts, counting skill) using the same methodology of this study (e.g., BEA, extended analysis, intervention phase, and follow-up). In fact, some third or fourth grade students were excluded from the study due to their lack of number sense and/or counting skills. Future research may examine the effects of potential interventions that address students' deficits in these prerequisite skills for calculation.

Additional Suggestions for Future Research

Future research should address the following issues to extend the present study.

Performance-Based Interventions

First, the study should be replicated with students who respond to performance-based interventions (e.g., contingency reinforcement, goal setting, and self-monitoring), a combination of skill-based and/or performance-based interventions (e.g., repeated practice plus contingency reinforcement), or intervention packages which include skill-based and/or performance-based intervention components (e.g., CCC; Skinner et al., 1986) to address the diverse deficits of students who are academically at-risk in mathematics (e.g., skill and performance deficits).

Effects of Empirically-Derived Interventions

Second, future research should further examine the effects of empirically-derived mathematics interventions on students' performance, satisfaction, and self-efficacy.

Single subject design is sensitive and useful to analyze students' idiosyncratic aspects that may predict their responsiveness to a certain intervention. Group design should be also used to identify common aspects of responders and nonresponders to an intervention as well as association among responsiveness, satisfaction, and self-efficacy. For example, a multivariate analysis of variance can be used to compare responsiveness (weekly slope), satisfaction, and self-efficacy between a group of students with and without SLD. Path analysis can be also used to examine the relationships among responsiveness, satisfaction, and self-efficacy. Based on the results of the study, it is assumed that responsiveness affects satisfaction and self-efficacy and/or vice versa. Other potential variables (e.g., grades, years of special education placement for students with SLD, etc.) may be also included in these analyses.

Influence of Social Validity and Self-Efficacy on Intervention Responsiveness

Third, future research may also examine how social validity (e.g., satisfaction with an intervention) and self-efficacy affect students' intervention responsiveness. For example, social validity affects a student's responsiveness by strengthening or weakening the association between his or her level of performance and responsiveness. This can be explained using terms mediator and moderator, which are commonly used in behavioral and pediatric research. A mediator is a variable through which an independent variable affects a dependent variable; a moderator is a variable which affects the association of independent and dependent variables, but is not affected by the independent variable or does not affect the dependent variable (Rose, Holmbeck, Coakley, & Franks, 2004). Researchers have demonstrated the mediator effects of self-efficacy on students' performance (e.g., Pajares & Kranzler, 1995; Stevens et al., 2004). It is assumed that self-efficacy and/or satisfaction is also a mediator for students' responsiveness (e.g., weekly slope, level increase, etc.) rather than performance at a single assessment opportunity. Whether social validity and/or self-efficacy are mediator and/or moderator may be analyzed using statistics procedures such as a series of regression analyses. Specifically, for a variable is to be a mediator, the following conditions must be met: an independent variable (e.g., initial levels of performance as measured by CBM) statistically significantly correlates with the variable (e.g., self-efficacy) in the first regression model; an independent variable also statistically significantly correlates with a dependent variable (e.g., a weekly slope) in the second regression model; and when the variable was entered into the second regression model, the significance between the independent and dependent variables statistically significantly decreases (i.e., partial mediation) or disappears (i.e., full mediation) (Baron & Kenny, 1986). Mediator and moderator effects of social validity may be analyzed in future research.

Reliability of BEA

Fourth, the reliability of BEA should be further examined in the area of mathematics. For example, it will be examined whether an intervention identified as most effective in BEA continues to enhance a student's performance in a subsequent extended analysis or intervention phase (i.e., test-retest reliability as for psychometric instruments). In this study, the intervention identified as most effective in BEA (i.e., the intervention that produced the highest DCPM) was identical to that identified across the BEA and extended analysis (i.e., the intervention that produced the highest mean DCPM) for four students. For Jacklyn, who demonstrated the equal highest DCPM under ICF twice and FI once in BEA, ICF was identified as most effective across the BEA and extended analysis. For Kaley, only four sessions in BEA yielded the most effective intervention. The results indicate that BEA may be a reliable method to identify an effective mathematics intervention. However, further empirical studies are warranted to examine this issue. Specifically, replication of BEA studies as well as meta-analyses of these studies (e.g., a meta-analysis in BEA in reading, Burns & Wagner, 2008) should be conducted in future research in mathematics. Additionally, in the present study, only Kaley demonstrated 20% increase under the same intervention twice in BEA. Future research may examine the aspects of students who are likely to demonstrate responsiveness that is a certain level higher under one intervention than the others.

Environmental Variables

Fifth, future research will examine environmental variables that affect a student's responsiveness, including a period of time for having been placed in special education, availability of individual instructions in class, the student's grade in mathematics, and areas where the student has been identified with SLD (e.g., only calculation or both calculation and reasoning). For example, the pervasive deficits in mathematics (e.g., SLD in both calculation and reasoning) of Kaley might have yielded her lower responsiveness than that of the other two students with SLD in calculation only. On the other hand, Tyson and Kathy had a strength in mathematics reasoning, and this might have yielded their higher responsiveness than Kaley. Fewer years of the SLD identification of Tyson (i.e., one year) might have yielded his higher responsiveness than the other two students with SLD (i.e., two to five years). Intensity of special education (e.g., the number of classes where a student attends in special education) might also have affected the students' responsiveness; Tyson and Kathy attended mathematics, reading, and English Language Arts classes in special education, whereas Kaley only attended a mathematics class in special education. This might indicate that intense education (e.g., special or resource education including small group, frequent individual instruction, immediate feedback, etc.) enhances the ability (e.g., memory), academic performance (e.g., intervention responsiveness), and/or motivation of students with academic deficits. Identification of variables that affect students' responsiveness may aid psychological practitioners (e.g., school psychologists) and educators (e.g., teachers) in early identifying students who are academically at-risk and referring these students for the RtI tier process.

Validity of the RtI Model

Finally, in school settings, the methodology used in this study (e.g., CBA, BEA, interventions, and follow-up assessment) can be used for examining a student's responsiveness in the last tier in the RtI model (e.g., Tier 3 in a typical RtI model, Level 4 in the Hartland Problem Solving Approach). As such, in future research, teachers (i.e., non-psychological professional) may identify the most effective intervention using BEA and implement the intervention for an additional period of time with the assistance of psychological professionals (e.g., instructional consultation provided by a school psychologist). In this case, the effects of an intervention as well as those of consultation on variables, including a student's performance, integrity of teacher implementation of the intervention, and the social validity of the intervention assessed with the student and teacher, will be examined. Whether teachers reliably determine if students are responders based on intervention data should be evaluated. This type of research may further examine the validity of the RtI model in SLD identification. The validity of the RtI model (e.g., discriminant validity as for psychometric instruments) may be demonstrated when the model identifies students who truly need special education (i.e., true positive) and those who can still be taught in general education with individual accommodations including individual interventions (i.e., true negative). The model should also decrease the number of students who are misidentified with SLD (i.e., false positive) or non-SLD (i.e., false negative). In the area of mathematics, further research is warranted to examine the effects of procedures used in responsiveness assessment (e.g., BEA, interventions, etc.) on students' performance in school settings, which in turn may enhance the validity of the RtI model.

Summary

The present study examined the effects of empirically-derived interventions on the responsiveness of students with and without SLD for mathematics calculation. Intervention responsiveness was compared for a pair of students with and without SLD who responded to the same intervention. Satisfaction and self-efficacy were also assessed with the students. The empirically-derived interventions were effective in enhancing students' responsiveness, satisfaction, and self-efficacy. However, more students without SLD were identified as responders than were those with SLD. All students had difficulty in generalizing their skills from instructional-level to grade-level materials.

This study can also be considered as a preliminary study to investigate the validity of the RtI model using a single subject design in school settings. The study partially demonstrated the validity of RtI such that implementation of empirically-derived interventions was effective to enhance the responsiveness of students with calculation deficits and to identify those who need further assessment (e.g., comprehensive evaluation) and individually-accommodated education (e.g., special education).

In summary, empirically-derived interventions may enhance responsiveness, satisfaction, and self-efficacy in students who are academically at-risk. However, further research is warranted to examine the validity of the RtI model in the area of mathematics.

REFERENCES

Arra, C. T., & Bahr, M. W. (2005). Teachers' and students' preferences for mathematics interventions: Implications for teacher acceptability in consultation. *Journal of Educational and Psychological Consultation, 16*(3), 157-174.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Belfiore, P. J., Skinner, C. H., & Ferkis, M. A. (1995). Effects of response and trial repetition on sight-word training for students with learning disabilities. *Journal of Applied Behavior Analysis, 28*(3), 347-348.

Bennett, K., & Cavanaugh, R. A. (1998). Effects of immediate self-correction, delayed self-correction, and no correction on the acquisition and maintenance of multiplication facts by a fourth-grade student with learning disabilities. *Journal of Applied Behavior Analysis, 31*(2), 303-306.

Bjorklund, D. F. (2005). *Children's thinking: Cognitive development and individual differences* (4^{th} ed.). Wadsworth/Thompson Learning: Belmont, CA.

Bonfiglio, C. M., Daly, E. J., III, Martens, B. K., Lin, L. H. R., & Corsaut, S. (2004). An experimental analysis of reading interventions: Generalization across instructional strategies, time, and passages. *Journal of Applied Behavior Analysis, 37*(1), 111-114.

Brown-Chidsey, R., & Steege, M. W. (2005). *Response to intervention: Principles and strategies for effective practice*. New York: Guilford.

Burns, M. K., & Wagner, D. (2008). Determining an effective intervention within a brief experimental analysis for reading: A meta-analytic review. *School Psychology Review, 37*(1), 126-136.

Caraway, K., Tucker, C. M., Reinke, W. M., & Hall, C. (2003). Self-efficacy, goal orientation, and fear of failure as predictors of school engagement in high school students. *Psychology in the Schools, 40*(4), 417-427.

Carnine, D., Jones, E. D., & Dixon, R. (1994). Mathematics: Educational tools for diverse learners. *School Psychology Review, 23*(3), 406-427.

Carson, P. M., & Eckert, T. L. (2003). An experimental analysis of mathematics instructional components: Examining the effects of student-selected versus empirically-selected interventions. *Journal of Behavioral Education, 12*(1), 35-54.

Cates, G. L. (2005). Effects of peer versus computer-assisted drill on mathematics response rates. *Psychology in the Schools, 42*(6), 637-646.

Cates, G. L., & Skinner, C. H. (2000). Getting remedial mathematics students to prefer homework with 20% and 40% more problems: An investigation of the strength of the interspersing procedure. *Psychology in the Schools, 37*(4), 339-347.

Codding, R. S., Shiyko, M., Russo, M., Birch, S., Fanning, E., & Jaspen, D. (2007). Comparing mathematics interventions: Does initial level of fluency predict intervention effectiveness? *Journal of School Psychology, 45*(6), 603-617.

Cooke, N. L., Guzaukas, R., Pressley, J. S., & Kerr, K. (1993). Effects of using a ratio of new items to review items during drill and practice: Three experiments. *Education and Treatment of Children*, 16(3), 213-234.

Cooke, N. L., & Reichard, S. M. (1996). The effects of different interspersal drill ratios on acquisition and generalization of multiplication and division facts. *Education and Treatment of Children, 19*(2), 124-142.

Cooter, K. S., & Cooter, R. B., Jr. (2004). One size doesn't fit all: Slow learners in the reading classroom. *The Reading Teacher, 57*(7), 680-684.

Daly, E. J., III, Andersen, M., Gortmaker, V., & Turner, A. (2006). Using experimental analysis to identify reading fluency interventions: Connecting the dots. *The Behavior Analyst Today, 7*(1), 133-150.

Daly, E. J., III, Lentz, F. E., Jr., & Boyer, J. (1996). The instructional hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly, 11*(4), 369-386.

Daly, E. J., III, & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis, 27*(3), 459-469.

Daly, E. J., III, Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis, 32*(1), 83-94.

Daly, E. J. III, Witt, J. C., Martens, B. M., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review, 26*(4), 554-574.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*(3), 219-232.

Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*(3), 358-374.

Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review, 33*(3), 429-443.

DuPaul, G. J., Ervin, R. A., Hook, C. L., & McGoey, K. E. (1998). Peer tutoring for children with attention deficit hyperactivity disorder: Effects on classroom behavior and academic performance. *Journal of Applied Behavior Analysis, 31*(4), 579-592.

Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools, 37*(5), 463-473.

Eckert, T. L., Ardoin, S. P., Daly, E. J., III, & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Behavior Analysis, 35*(3), 271-281.

Eckert, T. L., & Hintze, J. M. (2000). Behavioral conceptions and applications of acceptability: Issues related to service delivery and research methodology. *School Psychology Quarterly, 15*(2), 123-148.

Elliott, S. N. (1986). Children's ratings of the acceptability of classroom interventions for misbehavior: Findings and methodological considerations. *Journal of School Psychology, 24*(1), 23-35.

Elliott, S. N., Turco, T. L., & Gresham, F. M. (1987). Consumer's and client's pretreatment acceptability ratings of classroom group contingencies. *Journal of School Psychology, 25*(2), 145-153.

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*(2), 304-331.

Fletcher, J. M., Lyon, G. R., Barnes, M., Stuebing, K. K., Francis, D. J., Olson, R. K., Shaywitz, S. E., & Shaywitz, B. A. (2002). Classification of learning disabilities: An evidence-based evaluation. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 185-250). Mahwah, NJ: Lawrence Erlbaum.

Fuchs, D., Fuchs, L. S., & Fernstrom, P. (1993). A conservative approach to special education reform: Mainstreaming through transenvironmental programming and curriculum-based measurement. *American Educational Research Journal, 30*(1), 149-177.

Fuchs, L. S. (2005). Prevention research in mathematics: Improving outcomes, building identification models, and understanding disability. *Journal of Learning Disabilities, 38*(4), 350-352.

Fuchs, L. S., & Fuchs, D. (1997). Use of curriculum-based measurement in identifying students with disabilities. *Focus on Exceptional Children, 30*(3), 1-14, 16.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for recconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice, 13*(4), 204-219.

Fuchs, L. S., & Fuchs, D. (2001). Principles for the prevention and intervention of mathematics difficulties. *Learning Disabilities Research & Practice, 16*(2), 85-95.

Fuchs, L. S., & Fuchs, D. (2002). Curriculum-based measurement: Describing competence, enhancing outcomes, evaluating treatment effects, and identifying treatment nonresponders. *Peabody Journal of Education, 77*(2), 64-84.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children, 60*(6), 518-537.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review, 19*(1), 6-22.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal, 28*(3), 617-641.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*(1), 27-48.

Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities, 37*(4), 293-306.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*(2), 345-362.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4-15.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 293-304.

Gickling, E. E., & Thompson, V. P. (1985). A personal view of curriculum-based assessment. *Exceptional Children, 52*(3), 205-218.

Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*(1), 37-50.

Gresham, F. M. (2002). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 467-519). Mahwah, NJ: Lawrence Erlbaum.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*(2), 257-263.

Gresham, F. M., Gansle, K. A., Noell, G. H., Cohen, S., & Rosenblum, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980-1990. *School Psychology Review, 22*(2), 254-272.

Gresham, F. M., & Lopez, M. F. (1996). Social validation: A unifying concept for school-based consultation research and practice. *School Psychology Quarterly, 11*(3), 204-227.

Gresham, F. M., & Witt, J. C. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future direction. *School Psychology Quarterly, 12*(3), 249-267.

Hampton, N. Z., & Mason, E. (2003). Learning disabilities, gender, sources of efficacy, self-efficacy beliefs, and academic achievement in high school students. *Journal of School Psychology, 41*(2), 101-112.

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional technology: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23-40). Columbus, OH: Merrill.

Harter, S. (1985). *Manual for the Self-Perception Profile for Children: Revision of the Perceived Competence Scale for Children.* Unpublished manuscript, University of Denver, Colorado.

Hasselbring, T. S., Goin, L. I., & Bransford, J. D. (1988). Developing math automaticity in learning handicapped children: The role of computerized drill and practice. *Focus on Exceptional Children, 20*(6), 1-7.

Hawkins, J., Skinner, C. H., & Oliver, R. (2005). The effects of task demands and additive interspersal ratios on fifth-grade students' mathematics accuracy. *School Psychology Review, 34*(4), 543-555.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*(11), 963-974.

Henington, C., Kazmerski, J. S., Schuck, R., Ota, M., Watson, L. M., Adkins, H., & Dufrene, B. A. (2006, March). *Efficacy of brief academic intervention packages for academically at-risk children.* Symposium presented at the annual meeting of the National Association of School Psychologists, Anaheim, CA.

Individuals with Disabilities Education Improvement Act. (2004). Public Law 108-446 (20 U.S.C. 1400 *et seq*.).

Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: Further analysis of the brief assessment model for academic problems. *School Psychology Review, 31*(4), 554-568.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different forms of LD. *Journal of Learning Disabilities, 33*(6), 567-578.

Jordan, N. C., & Hanich, L. B. (2003). Characteristics of children with moderate mathematics deficiencies: A longitudinal perspective. *Learning Disabilities Research & Practice, 18*(4), 213-221.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74*(3), 834-850.

Kavale, K. A., & Reese, J. H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disabilities Quarterly, 15*(2), 74-94.

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7*(3), 164-177.

Kratochwill, T. R., & Shernoff, E. S. (2003). Evidence-based practice: Promoting evidence-based interventions in school psychology. *School Psychology Quarterly, 18*(4), 389-408.

Lackaye, T., Margalit, M., Ziv, O., & Ziman, T. (2006). Comparison of self-efficacy, mood, effort, and hope between students with learning disabilities and their non-LD-matched peers. *Learning Disabilities Research & Practice, 21*(2), 111-121.

Lee, M. J., & Tingstrom, D. H. (1994). A group math intervention: The modification of cover, copy, and compare for group application. *Psychology in the Schools, 31*(2), 133-145.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review, 31*(3), 313-327.

Logan, P., & Skinner, C. H. (1998). Improving students' perceptions of a mathematics assignment by increasing problem completion rates: Is problem completion a reinforcing event? *School Psychology Quarterly, 13*(4), 322-331.

Louisiana Department of Education (2004). *Division of student standards & assessments. Grade-Level Expectations – Mathematics.* Retrieved December 23, 2007 from http://www.doe.state.la.us/lde/saa/1841.html

Louisiana Department of Education (2005-2006). *LEAP/GEE annual report. Section I: The Louisiana Educational Assessment Program.* Retrieved December 23, 2007 from http://www.doe.state.la.us/lde/uploads/1703.pdf

Louisiana Department of Education (2006). *State special education data profile 2006.* Retrieved January 6, 2008, from http://www.doe.state.la.us/lde/uploads/10283.pdf

Louisiana Department of Education (2007, March). *District composite profile 2005-2006: Jefferson Parish.* Retrieved January 6, 2008, from http://www.louisianaschools.net/lde/pair/DCR0506/DCR026.pdf

Lovitt, T. C., & Hansen, C. L. (1976). The use of contingent skipping and drilling to improve oral reading and comprehension. *Journal of Learning Disabilities, 9*(8), 20-26.

Maccini, P., & Hughes, C. A. (1997). Mathematics interventions for adolescents with learning disabilities. *Learning Disabilities Research & Practice, 12*(3), 168-176.

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.

Marston, D., Muyskens, P., Lau, M., & Canter, A. (2003). Problem-solving model for decision making with high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research & Practice, 18*(3), 187-200.

Mercer, C. D., Jordan, L., Allsopp, D. H., & Mercer, A. R. (1996). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly, 19*(4), 217-232.

Mississippi Department of Education, Office of Special Education (October, 2002). *Policies & procedures: Definitions and eligibility criteria of disabilities.* Retrieved October 18, 2005, from http://www.mde.k12.ms.us/special_education/pdfs/New_Att_A.pdf

Muris, P., Meesters, C., & Fijen, P. (2003). The self-perception profile for children: Further evidence for its factor structure, reliability, and validity. *Personality and Individual Differences, 35*(8), 1791-1802.

No Child Left Behind Act. (2001). Public Law 107-15.

Noell, G. H., Freeland, J. T., Witt, J. C., & Gansle, K. A. (2001). Using brief assessments to identify effective interventions for individual students. *Journal of School Psychology, 39*(4), 335-355.

Pajares, F., & Kranzler, J. (1995). Self-efficacy and general mental ability in mathematics problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443.

Pellegrino, J. W., & Goldman, S. R. (1987). Information processing and elementary mathematics. *Journal of Learning Disabilities, 20*(1), 23-32, 57.

Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology, 95*(3), 589-603.

Reimers, T. M., Wacker, D. P., & Koeppl, G. (1987). Acceptability of behavioral interventions: A review of the literature. *School Psychology Review, 16*(2), 212-227.

Rhymer, K. N., Skinner, C. H., Henington, C., D'Reaux, R. A., & Sims, S. (1998). Effects of explicit timing on mathematics problem completion rates in African-American third-grade elementary students. *Journal of Applied Behavior Analysis, 31*(4), 673-677.

Rivera, D. P. (1997). Mathematics education and students with learning disabilities: Introduction to special series. *Journal of Learning Disabilities, 30*(1), 2-19, 68.

Robinson, S. L., & Skinner, C. H. (2002). Interspersing additional easier items to enhance mathematics performance on subsets requiring different task demands. *School Psychology Quarterly, 17*(2), 191-205.

Rose, B. M., Holmbeck, G. N., Coakley, R. M., & Franks, E. A. (2004). Mediator and moderator effects in developmental and behavioral pediatric research. *Developmental and Behavioral Pediatrics, 25*(1), 58-67.

Russell, R. L., & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematical difficulties. *Cognition and Instruction, 1*(2), 217-244.

Samuels, S. J. (1997). The method of repeated readings. *The Reading Teacher, 50*(5), 376-381. (Reprinted from *The Reading Teacher*, 1979, 32, 403-408.)

Schoolhouse Technologies, Inc. (2006). Mathematics Worksheet Factory (Version 3.0) [Computer software]. Seattle, WA: Schoolhouse Technologies, Inc.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3&4), 207-231.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis, 24*(2), 189-204.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

Shapiro, E. S. (1996). *Academic skills problems: Direct assessment and intervention* (2nd ed.). New York: Guilford.

Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention* (3rd ed.). New York: Guilford.

Shapiro, E. S., & Lentz, F. E., Jr. (1986). Behavioral assessment of academic skills. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 5, pp. 87-139). Hillsdale, NJ: Lawrence Erlbaum.

Shinn, M., & Marston, D. (1985). Differentiating mildly handicapped, low-achieving, and regular education students: A curriculum-based approach. *Remedial and Special Education, 6*(2), 31-38.

Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children.* New York: Guilford.

Shinn, M., R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: “Big ideas” and avoiding confusion. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement*. New York: Guilford.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*(8), 469-478, 486.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25*(10), 618-629.

Skinner, C. H. (1998). Preventing academic skills deficits. In T. S. Watson & F. M. Gresham (Eds.), *Handbook of child behavior therapy* (pp. 61-82). New York: Plenum.

Skinner, C. H. (2002). An empirical analysis of interspersal research evidence, implications, and applications of the discrete task completion hypothesis. *Journal of School Psychology, 40*(4), 347-368.

Skinner, C. H., Fletcher, P. A., & Henington, C. (1996). Increasing learning rates by increasing student response rates: A summary of research. *School Psychology Quarterly, 11*(4), 313-325.

Skinner, C. H., & Shapiro, E. S. (1989). A comparison of taped-words and drill interventions on reading fluency in adolescents with behavior disorders. *Education and Treatment of Children, 12*(2), 123-133.

Skinner, C. H., Shapiro, E. S., Turco, T. L., Cole, C. L., & Brown, D. K. (1992). A comparison of self- and peer-delivered immediate corrective feedback on multiplication performance. *Journal of School Psychology, 30*(2), 101-116.

Skinner, C. H., Turco, T. L., Beatty, K. L., & Rasavage, C. (1989). Cover, copy, and compare: A method for increasing multiplication performance. *School Psychology Review, 18*(3), 412-420.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407.

Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly, 17*(1), 47-77.

Sterling-Turner, H. E., Watson, T. S., Wildmon, M., Watkins, C., & Little, E. (2001). Investigating the relationship between training type and treatment integrity. *School Psychology Quarterly, 16*(1), 56-67.

Stevens, T., Olivarez, A., Jr., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *The Journal of Educational Research, 97*(4), 208-221.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367.

Tarnowski, K. J., & Simonian, S. J. (1992). Assessing treatment acceptability: The abbreviated acceptability rating profile. *Journal of Behavior Therapy & Experimental Psychiatry, 23*(2), 101-106.

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31*(4), 498-513.

Turco, T. L., & Elliott, S. N. (1986). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology, 24*(3), 277-283.

Van Houten, R., & Thompson, C. (1976). The effects of explicit timing on math performance. *Journal of Applied Behavior Analysis, 9*(2), 227-230.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69*(4), 391-409.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*(4), 601-638.

Von Brock, M., & Elliott, S. (1987). Influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology, 25*(2), 131-144.

Wildmon, M. E., Skinner, C. H., McCurdy, M., & Sims, S. (1999). Improving secondary students' perception on the "dreaded mathematics word problem assignment" by giving them more word problems. *Psychology in the Schools, 36*(4), 319-325.

Wildmon, M. E., Skinner, C. H., Watson, T. S., & Garrett, L. S. (2004). Enhancing assignment perceptions in students with mathematics learning disabilities by including more work: An extension of interspersal research. *School Psychology Quarterly, 19*(2), 106-120.

Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 251-288). Hillsdale, NJ: Lawrence Erlbaum.

Witt, J. C., & Martens, B. K. (1983). Assessing the acceptability of behavioral interventions used in classroom. *Psychology in the Schools, 20*(4), 510-517.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203-214.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock Johnson III Test of Achievement.* Itasca, IL: Riverside.

Woodward, J. (2004). Mathematics education in the United States: Past to present. *Journal of Learning Disabilities, 37*(1), 16-31.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strengths, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, 49(2), 156-167.

Ysseldyke, J. E., Algozzine, B., Shinn, M. R., & McGue, M. (1982). Similarities and differences between low achievers and students classified learning disabled. *Journal of Special Education, 16*(1), 73-85.

APPENDIX A

DEFINITION OF TERMS

Definition of Terms

Brief Experimental Analysis (BEA)

A single subject methodology in which potential interventions are implemented and their effects are compared.

Curriculum-Based Assessment (CBA)

Assessment conducted to identify the level of a student's academic performance.

Curriculum-Based Measurement (CBM)

Frustrational level

A level in which a student performs below an average range at the grade.

Instructional level

A level in which a student performs within an average range at the grade.

Mastery level

A level in which a student performs above an average range at the grade.

Digits correct per minute (DCPM)

The number of digits that were correctly stated on the worksheet.

Discrepancy Model

An identification model of specific learning disabilities (SLD) in which a certain discrepancy between test scores (e.g., one standard deviation of 15 points) yields identification of SLD. Intelligence quotient (IQ)-achievement or achievement-achievement discrepancy usually yields SLD identification in the discrepancy model.

Errors per minute (EPM)

The number of digits that were incorrect stated on the worksheet.

Folding-In Technique (FI)

An intervention in which a material including 70% to 85% of known problems and 15% to 30% of unknown problems is administered.

Follow-Up

Assessment conducted after the termination of an intervention to examine whether skills acquired through the intervention maintain.

Generalization

A skill acquired through an intervention is demonstrated under different stimuli such as time, place, or materials.

Immediate Corrective Feedback (ICF)

An intervention in which corrective feedback (e.g., error correction) is provided immediately after a student makes an error.

Repeated Practice (RP)

An intervention in which a student is required to complete the same material several times (e.g., three times) in order to reach a mastery level before an intervention datum is collected.

Response-to-Intervention (RtI) Model

An identification model of SLD in which a student's responsiveness to intervention is assessed and only a student who does not respond to research-based interventions is identified with SLD (i.e., nonresponder).

Self-Efficacy

An individual's perception of capability and competence on his/her current and future performance.

Skill Deficit

A type of deficit in which a student lacks skills to demonstrate academic performance rather than motivation.

Social Validity

Acceptability and social meaningfulness of intervention procedure and outcome.

Specific Learning Disabilities

A disability, identified under the Individuals with Disabilities Education Act of 2004, in which a student's academic level on a skill in a certain subject area (e.g., mathematics calculation) is lower than his/her grade level despite his/her sufficient intellectual ability.

Worksheets

CBA

Worksheets used to identify a student's instructional level based on the CBM criterion (Deno & Mirkin, 1977, cited in Shapiro & Lentz, 1986). CBA worksheets included multiple-skill problems for the assessed grade.

Instructional

Worksheets used to teach calculation skills during interventions and to assess maintenance of skills during follow-up. Instructional worksheets included single-skill problems at the student's instructional level..

Intervention monitoring

Worksheets used to examine progress during interventions. Intervention monitoring worksheets included single-skill problems at the student's instructional level.

Maintenance

Worksheets used to assess maintenance of the skill targeted during the intervention. Maintenance worksheets included single-skill problems at the student's instructional level.

Generalization

Worksheets used to assess generalization in calculation skills from the student's instructional-level materials to grade-level materials. Generalization worksheets included multiple-skill problems at the student's current grade.

APPENDIX B

PLACEMENT CRITERION FOR CURRICULUM-BASED MEASUREMENT IN MATHEMATICS

Placement Criterion for Curriculum-Based Measurement in Mathematics (Deno & Mirkin, 1977, cited in Shapiro & Lentz, 1986, p.124).

		Criterion	
Grade	Level	Median digits correct per minute	Median digits incorrect per minute
Grade 1-3	Frustrational	0 - 9	8+
	Instructional	10 - 19	3 - 7
	Mastery	20+	≤2
Grade 4-6	Frustrational	0 - 19	8+
	Instructional	20 - 39	3 - 7
	Mastery	40+	≤2

APPENDIX C

INSTITUTIONAL REVIEW BOARD APPROVAL LETTER

Mississippi State UNIVERSITY

August 21, 2008

Masanori Ota
PO Box 2057
Mississippi State, MS 39762

RE: IRB Study #07-105: Responsiveness of elementary-aged students to interventions, with and without specific learning disabilities, for mathematics calculation

Dear Mr. Ota:

The above referenced project was reviewed and approved via administrative review on 5/2/2007 in accordance with 45 CFR 46.101(b)(1). Continuing review is not necessary for this project. However, any modification to the project must be reviewed and approved by the IRB prior to implementation. Any failure to adhere to the approved protocol could result in suspension or termination of your project. The IRB reserves the right, at anytime during the project period, to observe you and the additional researchers on this project.

Please note that the MSU IRB is in the process of seeking accreditation for our human subjects protection program. As a result of these efforts, you will likely notice many changes in the IRB's policies and procedures in the coming months. These changes will be posted online at http://www.orc.msstate.edu/human/aahrpp.php.

Please refer to your IRB number (#07-105) when contacting our office regarding this application.

Thank you for your cooperation and good luck to you in conducting this research project. If you have questions or concerns, please contact Christine Williams at cwilliams@research.msstate.edu or call 662-325-5220.

Sincerely,

[For use with electronic submissions]

Christine Williams
IRB Compliance Administrator

cc: Carlen Henington

Office for Regulatory Compliance
P. O. Box 6223 • 70 Morgan Avenue • Mailstop 9563 • Mississippi State, MS 39762 • (662) 325-3294 • FAX (662) 325-8776

APPENDIX D

TEACHER INTERVIEW QUESTIONNAIRE

Teacher Interview Questionnaire

1. Name of student:

2. Age of student:

3. Grade of student:

4. Reason for nomination

5. Is the student receiving special education (yes or no)?

6. If so, what is the student's current exceptionality?

7. How long has the student been placed in special education?

8. What pre-referral interventions were implemented to improve his/her mathematics calculation performance?

9. What mathematics calculation interventions are currently being implemented with the student?

10. Parents' name:

11. Parents' phone number:

12. Parents' address:

APPENDIX E

SAMPLE MATHEMATICS WORKSHEET

Date: ____________________

Complete the Activity. DCPM = _______, EPM = _______

3 + 0	4 + 5	2 + 1	1 + 6	7 + 2
2 + 6	4 + 3	0 + 5	6 + 4	1 + 2
8 + 2	3 + 3	6 + 1	2 + 2	3 + 4
1 + 7	3 + 2	0 + 2	4 + 1	2 + 5
6 + 3	2 + 4	1 + 1	3 + 5	9 + 0
1 + 4	5 + 2	2 + 8	5 + 0	1 + 3

Single 30 1A-1-BEA1

APPENDIX F

STUDENT SATISFACTION SCALE FOR REPEATED PRACTICE

Student Satisfaction Questionnaire (RP)

Today you practiced several math worksheets. Please circle one that you agree with the most:

	True	*Somewhat True*	*Not True*
1. This intervention to improve my math skills was fair.	☺	😐	☹
2. My interventionist gave me enough time to practice math.	☺	😐	☹
3. This intervention is good one to use with other children.	☺	😐	☹
4. I like this intervention for my math skills.	☺	😐	☹
5. I think this intervention helps me do better in school.	☺	😐	☹

APPENDIX G

STUDENT SATISFACTION SCALE FOR

FOLDING-IN TECHNIQUE

Student Satisfaction Questionnaire (FI)

Today you practiced math problems using flashcards. Then, you completed a math worksheet for 1 minute. Please circle one that you agree with the most:

	True	*Somewhat True*	*Not True*
1. This intervention to improve my math skills was fair.	☺	😐	☹
2. My interventionist gave me enough time to practice math.	☺	😐	☹
3. This intervention is good one to use with other children.	☺	😐	☹
4. I like this intervention for my math skills.	☺	😐	☹
5. I think this intervention helps me do better in school.	☺	😐	☹

APPENDIX H

STUDENT SATISFACTION SCALE FOR IMMEDIATE CORRECTIVE FEEDBACK

Student Satisfaction Questionnaire (ICF)

Today you solved a math problem one by one. Your interventionist also checked your answer and corrected it if it is not correct. Then, you completed a math worksheet for 1 minute. Please circle one that you agree with the most:

	True	*Somewhat True*	*Not True*
1. This intervention to improve my math skills was fair.	☺	😐	☹
2. My interventionist gave me enough time to practice math.	☺	😐	☹
3. This intervention is good one to use with other children.	☺	😐	☹
4. I like this intervention for my math skills.	☺	😐	☹
5. I think this intervention helps me do better in school.	☺	😐	☹

APPENDIX I

STUDENT SELF-EFFICACY SCALE

Student Self-Efficacy Questionnaire

Please circle one that you agree with the most:

	True	*Somewhat True*	*Not True*
1. I feel I am good at math.	☺	😐	☹
2. I feel I am as smart as my classmates.	☺	😐	☹
3. I feel I can do math problems quickly.	☺	😐	☹
4. I feel I can memorize math problems easily.	☺	😐	☹
5. I feel I can figure out the answers almost always.	☺	😐	☹

APPENDIX J

PROCEDURE FOR SCORING DIGITS CORRECT PER MINUTE AND ERRORS PER MINUTE

Procedure for Scoring Digits Correct Per Minute (DCPM) and Errors Per Minute (EPM)

1. ______ Check if each digit is correct using an answer key
2. ______ A number indicating carrying or borrowing written above the problem by a student is not counted as DCPM or EPM
3. ______ Count a total number of DCPM
4. ______ Count a total number of EPM
5. ______ The following cases are not considered errors
    a. 0 in the highest digit; for example, if a student writes 06 for a problem of 12-6, 0 is not considered as an error
6. ______ The following cases are considered as errors
    a. A blank digit is considered as error; for example, if a student writes 20 for a problem of 140-20, a missing 1 in the 100 column is considered as an error
    b. When a student skips a problem, all digits in the answer for the problem are considered errors; for example, if a student skips a problem of 20+40, this indicates 2 errors
    c. If 1 minute elapses when a student is solving a problem and has not completed the problem, missing digits are considered as errors; for example, if 1 minute elapsed when a student has written 4 for a problem of 24-10, a missing 1 in the 10 column is counted as an error

APPENDIX K

PROCEDURAL INTEGRITY CHECKLIST FOR

PRE-TREATMENT ASSESSMENT

Procedural Integrity Checklist

Pre-Treatment Assessment

Please mark each step that you completed.

________ Step 1: Give an instructional worksheet to the student.

________ Step 2: Tell the student, "You have 1 minute to complete this worksheet. Solve problems across lines. Work as quickly as you can. Begin."

________ Step 3: Time for 1 minute.

________ Step 4: Tell the student, "Stop," after 1 minute.

________ Step 5: Calculate DCPM.

________ Step 6: Repeat the Steps 1 to 5 two more times.

________ Step 7: Calculate median DCPM.

________ Step 8: If the median DCPM is in the instructional level, stop here; if in the mastery level, repeat the Steps 1 to 5 three times using a one grade above level materials; if in the frustrational level, repeat the steps using a one grade below level materials.

________ Step 9: Continue until instructional level is obtained.

APPENDIX L

TREATMENT INTEGRITY CHECKLIST FOR

REPEATED PRACTICE

Treatment Integrity Checklist

Repeated Practice

Please mark each step that you completed.

________ Step 1: Give an instructional worksheet to the student.

________ Step 2: Tell the student, "Solve problems across lines. Work as quickly as you can. Let me know after you are finished. Begin."

________ Step 3: Give the same instructional worksheet to the student.

________ Step 4: Repeat Steps 1 to 4 using the same instructional worksheet.

________ Step 5: Repeat Steps 1 to 3 using the same instructional worksheet.

________ Step 6: Give another instructional worksheet including the same problems as in the practice worksheet with different order

________ Step 7: Tell the student, "You have 1 minute to complete this worksheet. Solve problems across lines. Work as quickly as you can. Begin."

________ Step 8: Time for 1 minute.

________ Step 9: Tell the student, "Stop," after 1 minute.

________ Step 10: Calculate DCPM and EPM and write them on the worksheet.

APPENDIX M

TREATMENT INTEGRITY CHECKLIST FOR

FOLDING-IN TECHNIQUE

Treatment Integrity Checklist

Folding-In Technique

Please mark each step that you completed.

________ Step 1: Show each of the ten flashcards containing seven (70%) known and three (30%) unknown facts.

________ Step 2: If the student responds to a fact correctly within 3 seconds, present the next flashcard; if the student responds incorrectly or did not respond within 3 seconds, tell the correct answer to the student without explaining the calculation procedure and present the next flashcard.

________ Step 3: Continue Step 2 until all ten flashcards are presented.

________ Step 4: Shuffle all flashcards.

________ Step 5: Repeat Steps 2, 3, and 4 using the same ten flashcards.

________ Step 6: Repeat Steps 2 and 3 using the same ten flashcards

________ Step 6: Give an instructional worksheet including the same problems as in the folding-in technique with different order

________ Step 7: Tell the student, “You have 1 minute to complete this worksheet. Solve problems across lines. Work as quickly as you can. Begin.”

________ Step 8: Time for 1 minute.

________ Step 9: Tell the student, “Stop,” after 1 minute.

________ Step 10: Calculate DCPM and EPM and write them on the worksheet.

APPENDIX N

TREATMENT INTEGRITY CHECKLIST FOR

IMMEDIATE CORRECTIVE FEEDBACK

Treatment Integrity Checklist

Immediate Corrective Feedback

Please mark each step that you completed.

________ Step 1: Give an instructional worksheet to the student.

________ Step 2: Tell the student, “Solve each problem.”

________ Step 3: Each time a student completes a problem, check the answer.

________ Step 4: If the student responded correctly, say “Right!” However, if the student responded incorrectly, tell the student the correct answer, briefly explain the calculation procedure, and have the student solve the next problem.

________ Step 5: Repeat Steps 3 and 4 until the student completes the worksheet.

________ Step 6: Give another instructional worksheet containing the same problems as in the practice worksheet in a different order

________ Step 7: Tell the student, “You have 1 minute to complete this worksheet. Solve problems across lines. Work as quickly as you can. Begin.”

________ Step 8: Time for 1 minute.

________ Step 9: Tell the student, “Stop,” after 1 minute.

________ Step 10: Calculate DCPM and EPM and write them on the worksheet.

APPENDIX O

PROCEDURAL INTEGRITY CHECKLIST FOR

FOLLOW-UP

Procedural Integrity Checklist

Follow-Up

Please mark each step that you completed.

________ Step 1: Give an instructional worksheet to the student.

________ Step 2: Tell the student, “You have 1 minute to complete this worksheet. Solve problems across lines. Work as quickly as you can. Begin.”

________ Step 3: Time for 1 minute.

________ Step 4: Tell the student, “Stop,” after 1 minute.

________ Step 5: Calculate the DCPM and EPM and write them on the worksheet.

________ Step 6: Repeat Steps 1 to 5 two more times.

________ Step 7: Calculate median DCPM and EPM.

________ Step 8: Give a generalization worksheet to the student.

________ Step 9: Repeat Steps from 2 to 5.

APPENDIX P

CURRICULUM VITA

CURRICULUM VITA
MASANORI OTA

Masanori Ota, M.S.
Address: P.O. Box 2057
Mississippi State, MS 39762
E-mail: mo36@msstate.edu
Phone: (662) 722-8051

Licensed Psychometrist	State of Mississippi AA, November 2006 – present (#194903)

SECTION I: EDUCATION

Ph.D. Educational Psychology (School Psychology)	Mississippi State University – APA/NASP Accredited Program Mississippi State, MS – Anticipated December 2008 Dissertation Title: Responsiveness of elementary-aged students, with and without specific learning disabilities, to interventions for mathematics calculation.
M.S. Educational Psychology (Psychometry)	Mississippi State University – APA/NASP Accredited Program Mississippi State, MS – August 2005
M.S. Mathematics	University of Tsukuba, Ibaraki, Japan – March 1999
B.Ed. Mathematics	Tokyo Gakugei University, Tokyo, Japan – March 1997

SECTION II: RELEVANT APPLIED EXPERIENCES

August 2007 – July 2008	Pre-doctoral school psychology internship Louisiana School Psychology Internship Consortium, New Orleans, LA Currently working as a school psychologist intern in Jefferson Parish Public Schools, LA. Duties include conducting assessment, developing intervention, provide consultation to teachers/parents, and conducting psychological evaluations.

August 2006 – May 2007	Behavior Specialist in Starkville School District, MS. Supervisors: Ms. Sandi George (onsite), Dr. Kristin Johnson-Gros (university). Conducted functional behavioral assessment, developed behavioral intervention plans, and provided consultation to teachers and administrators (20 hours per week).
May 2006 – August 2006	Development of *Reading to Read* intervention probes. Supervisor: Dr. Kristin Johnson-Gros. Developed reading probes (e.g., comprehension and maze probes) of the reading intervention packet, *Reading to Read*.
January 2006 – present	Curriculum-based measurement (CBM) prediction to statewide achievement tests. Supervisor: Dr. Kristin Johnson-Gros. Assisted Dr. Johnson-Gros' research on CBM prediction. Conducted statistical analyses (e.g., correlation, regression, etc.).
January 2006 – May 2006	Teaching Assistant of Ms. Annalisa Ebanks. Mississippi State University, MS. Taught graduate students who majored in educational psychology and related areas basic maneuvers of SPSS (e.g., central tendency, *t*-test, ANOVA) in EPY 6214 lab. Scored students' homework.
January 2006 – May 2006	School Psychology Consultation Practicum (300 hours) West Point School District, MS. Supervisor: Dr. Kristin Johnson-Gros. Conducted behavioral and academic assessment, developed academic and behavioral interventions, and provided consultation to general education teachers.
August 2005 – May 2006	Teaching Assistant of Dr. Anastasia Elder. Mississippi State University, MS. Taught undergraduate students who majored in educational psychology and related areas basic maneuvers of SPSS (e.g., central tendency, correlation, *t*-tests) in EPY 4214 lab.

August 2005	School Readiness Developmental Screening Test Starkville, MS. Supervisor: Dr. Carlen Henington. Conducted a developmental screening test, Developmental Indicators for the Assessment of Learning – Revised (DIAL-R) with approximately 5 children aged from 3-0 through 6-11.
July 2005 – August 2005	Summer Academic Clinic (100 hours). Mississippi State University, MS. Supervisor: Drs. Carlen Henington, R. Anthony Doggett, and Brad Dufrene. Implemented an intervention package, *Math to Mastery*, with approximately 6 elementary school students with skill/performance deficits in mathematics calculation.
January 2005 – May 2005	School Psychology Assessment Practicum (300 hours). Starkville School District, MS. Supervisor: Ms. Celeste King (onsite), Dr. Harrison Kane (university). Conducted intelligence tests (e.g., WISC-III & IV, K-ABC, Leiter-R) and achievement tests (e.g., WIAT-II, WJ-III ACH, K-TEA).
August 2004 – May 2005	Teaching Assistant of Dr. Anastasia Elder. Mississippi State University, MS. Taught undergraduate students who majored in educational psychology and related areas basic maneuvers of SPSS (e.g., central tendency, correlation, *t*-tests) in EPY 4214 lab.
August 2004	School Readiness Developmental Screening Test. Starkville, MS. Supervisor: Dr. Carlen Henington. Conducted a developmental screening test, Developmental Indicators for the Assessment of Learning - Revised (DIAL-R) with approximately 5 children aged from 3-0 through 6-11.
July 2004 – August 2004	Summer Academic Clinic (100 hours). Mississippi State University, MS. Supervisor: Dr. Carlen Henington. Implemented mathematics interventions with approximately 5 elementary school students with skill/performance deficits in mathematics calculation.

March 2004 – April 2004	Conducted a study in which reading interventions (e.g., repeated reading and listening passage previewing) were implemented with an elementary school child (first grade) with autism with another graduate student (10 hours). Supervisor: Dr. Sandy Devlin.
August 2003 – May 2004	Teaching Assistant of Dr. Susan Fascio-Vereen. Mississippi State University, MS. Assisted Dr. Fascio-Vereen's research on reading interventions conducted in school settings (e.g., repeated reading and listening previewing) and helped teaching courses (e.g., exam administration).
April 1999 – March 2003	Full-time Mathematics Teacher, Hosei Daini High School Kanagawa, Japan. Taught mathematics courses (e.g., high school algebra, geometry, and statistics). Guided approximately 40 students as a homeroom teacher each year.
April 1998 – March 1999	Part-time Mathematics Teacher, Tsukuba International University-Chiyoda High School, Ibaraki, Japan. Supervisor: Mr. Shin Terada. Taught mathematics (e.g., high school algebra).
April 1998 – March 1999	Teaching Assistant of Dr. Takao Hoshina. University of Tsukuba, Ibaraki, Japan. Assisted Dr. Hoshina' research on general topology (e.g., preparing conference handouts).

SECTION III: PUBLICATIONS

Kuhn, L., Watson, L., Ota, M., Cole, M., & Johnson-Gros, K.N. (in press). Effects of a brief experimental analysis and reading intervention on fluency, comprehension, and high word overlap. *Journal of Evidence-Based Practices for Schools.*

Ota, M. (2005). Review of the Brief Symptom Inventory (BSI). In B. T. Erford (Ed.), *The counselor's guide to personality, clinical, and behavioral assessment* (pp. 37-40). Lahaska, PA: Lahaska Press.

Sheperis, C. J., Doggett, R. A., Ota, M., Erford, B. T., & Salisbury, C. (2007). Behavioral assessment. In B. T. Erford (Ed.), *Assessment for counselors* (pp. 303-318). Boston, MA: Hughton Mifflin Company.

SECTION IV: PRESENTATIONS

<u>2007</u>

Ota, M. (2007, September). A/BIT Model as preventive and proactive services. Workshop presented in the faculty meeting at Cherbonnier Elementary School in Jefferson Parish Public Schools.

<u>2006</u>

Watson, L., Ota, M., Kuhn, L., Cole, M., Miller, M., & Johnson-Gros, K. (2006, November). *Brief experimental analysis and reading intervention with consultation.* Presented at the annual meeting of the Mid South Educational Research Association, Birmingham, AL.

Ota, M., Smith, S., & Kazmerski, J. (2006). *Functional behavior assessment.* Workshop presented at the administrator meeting in Starkville School District.

Henington, C., Kazmersk, J., Campbell, K. W., Schuck, R., Ota, M., Watson, L. M., Smith, S., Adkins, H., Rye, D. A., & Dufrene, B. A. (2006, April). *Efficacy of brief academic intervention packages for academically at-risk children.* Symposium presented at the annual meeting of the National Association of School Psychologists, Anaheim, CA.

Kazmerski, J., Ota, M., Schuck, R., Campbell, K., Doggett, R. A., & Henington, C. (2006, May). *Math-to-Mastery: An examination of a mathematics intervention package to increase math fluency.* Poster presentation at the annual meeting of the Association of Behavior Analysis, Atlanta, GA.

Kuhn, L., Watson, L., Cole, M., Ota, M., & Johnson-Gros, K. (2006, February). Effects of brief experimental analysis to identify a reading intervention. Symposium presented at the annual conference of Mississippi Association for Psychology in the Schools, Jackson, MS.

<u>2005</u>

Henington, C., Jones, C., Graves, S., Bowers, V., Baker, J., Mong, M., Smith, C.S., Ota, M., Baylot, L., & Emens, B. (2005, April). *Brief interventions for mathematics, reading, and writing with elementary children.* Symposium presented at the annual meeting of the National Association of School Psychologists, Atlanta, GA.

Mong, M., Ota, M., Campbell, C., Dufrene, B.A., & Doggett, R.A. (2005, April). *Differential effects of compliant behavior displayed by a preschooler during clinic-based brief experimental analysis and intervention sessions when conducted by different adults*. Symposium presented at the annual conference of Mississippi Association for Psychology in the Schools, Jackson, MS.

Ota, M., Mong, M., Davis, C., & Henington, C. (2005, April). *Curriculum based math interventions with elementary and middle school students*. Symposium presented at the annual conference of Mississippi Association for Psychology in the Schools, Jackson, MS.

Ota, M., Kuhn, L., Bodkin, A., Dollar, J., Campbell, C., Dufrene, B., & Kane, H. (2005, April). *Response to intervention: Intervention results for students in learning disability categories*. Symposium presented at the annual conference of Mississippi Association for Psychology in the Schools, Jackson, MS.

2004

Henington, C., Mong, M., Bodkin, A., Ota, M., Kuhn, L., Jones, C., & Smith, C.S. (2004, November). *Effectiveness of academic interventions implemented in a brief summer academic clinic*. Symposium presented at the annual meeting of the Mid South Educational Research Association, Gatlinburg, TN.

Jones, C., Bowers, V. L., & Ota, M. (2004, October). *Cultural diversity on the rise: Increasing awareness in the schools*. Presented at Mid-South Regional Conference on Psychology in the Schools, Tunica, MS.

SECTION V: TRAINING

2007

September 2007	ABIT training in Jefferson Parish Public Schools University of New Orleans, LA. Training regarding the administration procedures of ABIT screening (DIBELS and CBM) and the implementation of the PAM Model.

2006

August 2006	Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and Test of Early Numeracy (TEN) training. Dr. Kristin Johnson-Gros. Mississippi State University, MS. Training regarding the administration procedures of the DIBELS and TEN.

2005

November 2005	Functional behavioral assessment (FBA) training. Dr. Kristin Johnson-Gros. Training regarding the rationale and research background of FBA and procedures of FBA.

September 2005	Hurricane Katrina Response. Dr. Phil Lazarus. Oxford, MS. Training regarding crisis intervention and consultation that address the aftermath of Hurricane Katrina.
October 2005	Hurricane Katrina and Crisis Intervention. Dr. Joe Olmi. Mississippi State University, MS. Training regarding crisis intervention and consultation that address the aftermath of Hurricane Katrina implemented in southern Mississippi areas.
<u>2004</u>	
August 2004	Developmental Indicators for the Assessment of Learning – Revised (DIAL-R) test administration training. Dr. Carlen Henington. Mississippi State University, MS. Training regarding the test administration procedure of the DIAL-R.

SECTION VI: MEMBERSHIP IN PROFESSIONAL ORGANIZATIONS

Mississippi Association for Psychology in the Schools (MAPS):
    Spring 2004 – December 2006.
National Association of School Psychologists (NASP): Spring 2004 – present.
American Psychological Association (APA)
    Division 16 (student affiliation membership): Spring 2004 – present.
Association for Behavior Analysis: Spring 2006 – present.