

RESPONSE TO INTERVENTION:
INCORPORATION OF AN INCREASING
INTENSITY DESIGN TO
IMPROVE MATHEMATICS
FLUENCY

By

MICHELLE ELIZABETH ATKINS

Bachelor of Arts
University of St. Thomas
Houston, Texas
2000

Master of Arts
University of Houston
Houston, Texas
2003

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2008

RESPONSE TO INTERVENTION:
INCORPORATION OF AN INCREASING
INTENSITY DESIGN TO
IMPROVE MATHEMATICS
FLUENCY

Dissertation Approved:

Dr. Gary Duhon

Dissertation Adviser

Dr. Terry Stinnett

Dr. Eric Mesmer

Dr. Heidi Mesmer

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGEMENTS

It truly takes a village to raise a dissertation. I would like to express my appreciation to the staff and administration of Perry Independent School District for allowing me to work with their students and helping me to gather my data in a timely manner. I wish to thank my committee members for their endless support with this project. I would like to thank my chair, Terry Stinnett, Ph.D. for his encouragement, feedback, and superb editorial skills. I would also like to thank Eric Mesmer, Ph.D. for his input, support, and the substantial amount of knowledge he provided throughout this process. I am also grateful to Heidi Mesmer, Ph.D. for providing alternative perspectives and methodological guidance, both which were invaluable to this project. I am most indebted to my advisor, Gary Duhon, Ph.D., who gave so willingly of his time, experience, and knowledge during my graduate career at Oklahoma State University, especially during the completion of this project; I'm appreciative of the time he spent listening to my ideas and for the encouragement to follow my interests. I would also like to thank my family and friends for their endless support during this time. I am most grateful for my research assistants, who without their help this dissertation would not have been possible: Shilah Scherweit, Shannon Beason, Kim Wiechmann, Cari Fellers, Heather Hale, Angela Hamlin, Valexia Hampton, Lindsey Thompson, Tiffany Truitt, Cassie Wong, and Nina Ellis Hervey. Lastly I

would like to thank my fiancé, Nick, for his endless patience throughout my entire graduate career. His encouragement during a time when I decided to change career paths is invaluable. Without his support this degree would have not been possible.

TABLE OF CONTENTS

| Chapter | Page |
|---|------|
| I. INTRODUCTION | 1 |
| Response to Intervention Models..... | 2 |
| Intervention Response Options in RTI | 4 |
| Intervention Intensity as Indicator of Response to Intervention | 6 |
| Math Disabilities and Response to Intervention | 7 |
| Empirically Supported Math Interventions..... | 8 |
| Rationale..... | 9 |
| II. REVIEW OF LITERATURE..... | 13 |
| History of Learning Disabilities Assessment..... | 13 |
| The Discrepancy Model..... | 13 |
| Response to Intervention | 15 |
| The Original Treatment Validity Model | 16 |
| Using Curriculum Based Measurement..... | 19 |
| Response to Intervention Models..... | 23 |
| Problem Solving | 23 |
| Standard Protocol Approach | 26 |
| Comparison of Problem Solving and Standard Protocol | 31 |
| Intervention Response Options in RTI | 33 |
| Measurement Timing..... | 33 |
| Final Status Model..... | 33 |
| Growth Model..... | 34 |
| Dual Discrepancy Model | 35 |
| Standards For Designating Response..... | 37 |
| Nature of Intervention..... | 38 |
| Response Intervention Models by Academic Area..... | 41 |
| Math Interventions..... | 42 |
| Reinforcement..... | 42 |
| Goal Setting | 43 |
| Rationale..... | 44 |
| Research Questions and Hypotheses..... | 46 |

| Chapter | Page |
|--|------|
| III. METHODS..... | 48 |
| Participants and Setting | 48 |
| Materials..... | 49 |
| Single-Skill Probe..... | 49 |
| Reinforcements..... | 49 |
| Dependent Variable | 50 |
| Fluency Intervention..... | 51 |
| Overview of Procedures and Experimental Design | 51 |
| Pre-Intervention Phase..... | 52 |
| Baseline | 52 |
| Accuracy vs. Fluency | 53 |
| Intervention Phase One..... | 54 |
| Proficiency Group Identification..... | 55 |
| Intervention Phase Two..... | 56 |
| Baseline | 56 |
| Intensity of Two Sessions Daily..... | 57 |
| Intensity of Four Sessions Daily | 57 |
| Baseline | 52 |
| Accuracy vs. Fluency | 53 |
| Post Intervention Phase | 58 |
| Reliability Data | 58 |
| Planned Statistical Analyses | 59 |
| IV. RESULTS..... | 60 |
| Intervention Phase One | 60 |
| Proficiency Group Identification..... | 63 |
| Intervention Phase Two | 67 |
| Baseline | 67 |
| Intensity of Two Sessions Daily..... | 67 |
| Intensity of Four Sessions Daily | 68 |
| Baseline | 67 |
| Post Intervention Phase | 68 |
| V. DISCUSSION..... | 70 |
| Implications for Practice | 72 |
| Limitations and Directions for Future Research | 74 |
| REFERENCES..... | 77 |

| Chapter | Page |
|---------------------------------------|------|
| APPENDICES | 85 |
| Appendix A: Research Prospectus..... | 85 |
| Appendix B: Single Skill Probe | 88 |
| Appendix C: IRB Approval Letter | 89 |

LIST OF TABLES

| Table | Page |
|--|------|
| 1. Descriptive Statistics of the Final Sample..... | 63 |
| 2. Descriptive Statistics for Response Groups..... | 63 |
| 3. Analysis of Variance of Fluency Slopes..... | 65 |
| 4. Games-Howell Post Hoc Results..... | 66 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1. Distribution of Fluency Slopes of the Total Sample | 61 |
| 2. Distribution of Fluency Slopes of the Final Sample | 62 |
| 3. Mean Slopes of Response Groups..... | 64 |
| 4. Error Bar Chart of Slope by Response Group | 66 |
| 5. Multiple Baseline Design Across Subjects..... | 69 |

NOMENCLATURE

| | |
|--------|---|
| ADD | Auditory Discrimination in Depth program |
| ARR | Average Rate Responder |
| ANOVA | analysis of variance |
| CBA | curriculum-based assessment |
| CBM | curriculum-based measurement |
| CBM-DD | curriculum-based measurement, dual discrepancy |
| CTOPP | Comprehensive Test of Phonological Processing |
| dc | digits correct |
| dcpm | digits correct per minute |
| dcp2m | digits correct per 2 minutes |
| DIBELS | Dynamic Indicators of Basic Early Literacy Skills |
| EP | Embedded Phonics program |
| F | F-statistic, measure of correlation |
| FDD | Frequently Dually Discrepant |
| GLM | General Linear Model |
| GORT-4 | Gray Oral Reading Tests, Fourth Edition |
| HOSTS | Help One Student to Succeed program |
| HRR | High Rate Responder |
| IRB | Institutional Review Board |

| | |
|--------|--|
| IDD | Infrequently Dually Discrepant |
| IDEA | Individuals with Disabilities Education Act |
| IQ | intelligence quotient (mean = 100, SD = 15) |
| IQ-DS | regression based IQ-reading achievement discrepancy |
| LA | low achievement |
| LD | learning disabilities |
| LRR | Low Rate Responder |
| n | number |
| NDD | Never Dually Discrepant |
| NRC | National Research Council |
| ORF | Oral Reading Fluency |
| p | probability, usually of error (p = .05 means 5% chance of error) |
| PASS | Priority Academic Student Skill |
| RAN | rapid automatized naming |
| RTI | Response to Intervention |
| SD | standard deviation |
| SPSS | Statistical Package for the Social Sciences |
| TORF | Test of Oral Reading Fluency |
| TPRI | Texas Primary Reading Inventory |
| WJ-R | Woodcock Johnson – Revised |
| WRMT-R | Woodcock Reading Mastery Tests – Revised |

CHAPTER I

INTRODUCTION

Research regarding learning disabilities (LD) has surfaced from social and educational necessity as a response to (1) the need to understand differences in children and adults who were identified as having specific deficits in spoken or written language while maintaining normative intellectual functioning and (2) to provide remedial services to these individuals (Lyon, 1996). Diagnostic practices in this field, however, are believed to have emerged from practice, law, and policy rather than from empirical support (Gresham, VanDerHeyden, & Witt, 2005; Lyon, 1996). Practices in learning disability assessment have been viewed with scrutiny due to the recent reauthorization of the Individuals With Disabilities Education Act (IDEA). The discrepancy model, used for diagnosis of learning disabilities over the last three decades, has been highly criticized and researchers have presented several views concerning the need for change (Algozzine & Ysseldyke, 1986). A criticism surrounding discrepancy models is that they fail to differentiate low achievement from learning disabilities, overidentify students, are not implemented consistently, lack treatment validity, and delay the student's access to intervention since many approaches are based on a "wait to fail" model (Aaron, 1997; Dean & Burns, 2002; Gresham, 2001; Proctor & Prevatt, 2003). Changes in the current assessment approaches are

relevant to researchers, practitioners, teachers, and parents because these modifications may reduce the overidentification of students with learning disabilities and allow for effective interventions for those students in need of special education services.

Response to intervention (RTI) is defined as a change in performance or behavior due to intervention (Fuchs, Mock, Morgan, & Young, 2003; Gresham, 1991; Gresham, 2001). The approach does have a discrepancy base, but the discrepancy is between pre- and post-intervention levels instead of ability and achievement scores. This model deals with a *lack* of discrepancy since the goal of intervention is to produce a difference in scores. According to the response to intervention approach, those who are nonresponsive to treatment (even after it has been extended, intensified, or changed) might have a learning disability or other disability.

Response to Intervention Models

Currently, there are two fundamental frameworks for RTI models: (1) the problem solving approach and (2) the standard protocol approach (Fuchs et al., 2003, Gresham, VanDerHeyden, & Witt, 2005). The problem solving approach derives from the behavioral consultation literature which was first detailed by Bergan (1977) and Bergan and Kratochwill (1990) (Fuchs et al., 2003; Gresham, VanDerHeyden, & Witt, 2005). It is an inductive model that uses a four step process involving problem identification, problem analysis, plan implementation, and problem evaluation to guide the intervention process (Telzrow, McNamara, & Hollinger, 2000). Throughout all phases data are collected to evaluate the

student's response to the intervention which results in an intervention specifically tailored to the student's needs. The problem solving model is favored by practitioners, but has limited empirical support regarding its use as an approach to determine eligibility (Fuchs et al., 2003).

The standard protocol approach uses the same empirically validated intervention across all students experiencing difficulties in a particular academic domain (e.g., addition) (Fuchs et al., 2003, Gresham, VanDerHeyden, & Witt, 2005). This approach is found to be helpful in differentiating between those students who may have problems due to actual deficits rather than problems due to lack of practice or exposure to instruction (Fuchs et al., 2003; Torgesen et al., 2001; Vellutino et al., 1996). The standard protocol approach assumes if a student responds to an intensive treatment trial then he/she does not have a disability, the problem can be remediated, and he/she can receive instruction in the general education classroom (Fuchs, Fuchs, & Compton, 2004).

The reasoning and structure behind both of these models produce advantages and disadvantages for each. The problem solving approach is very sensitive to individual differences. This in turn could produce very effective outcomes for those students. However, because this process involves working with individual students it is more difficult to implement. Procedures requiring more time and resources may not be readily accepted. Ysseldyke (2001) states "Change is difficult and more political than data-based...and while change is difficult, change requiring extra work is next to impossible" (pg. 300). Though this model may be able to produce beneficial outcomes for students, more research

streamlining the approach and documenting its effectiveness is required. On the other hand, the standard protocol approach has greater quality control. This approach could produce outcomes for large groups of students across varying levels of performance. However, in this model students typically do not receive individualized instruction. The students who need intervention the most may not be able to receive the intensity of the intervention they require to maximize their progress. Neither model has produced research yet as to how they can be adopted on a large scale (Fuchs, et al., 2003).

Intervention Response Options in RTI

Several approaches to RTI have been researched in the last decade. These methods differ in the timing of measurement (final status, growth, or dual-discrepancy), standards for designating responsiveness, and the nature of intervention (intensive versus general education) (Fuchs, 2003). Each RTI model has its own implications and issues.

There are three types of timing procedures documented in the RTI literature: final status, growth, and dual discrepancy. *Final status* involves testing the student at the end of the intervention period, with the student's responses based on his or her post-intervention outcome (Torgesen et al., 2001). *Growth* models measure the student's progress periodically throughout the intervention in order to ascertain how much learning took place (Vellutino et al., 1996). They seek to produce a discernable difference between pre-intervention and post-intervention levels of performance (Gresham, VanDerHeyden, & Witt 2005). *Dual discrepancy* models determine if the student's level and slope are

significantly lower than that of his/her peers (Case, Speece, & Malloy, 2003; Fuchs & Fuchs, 1998).

No matter which model is used, a second evaluative standard must also be applied to the measurement to create a cutpoint for differentiating responders from nonresponders (Fuchs, 2003). This criterion can be gauged from viewing the full distribution of student functioning, which is referred to as the *Normative* approach. Several have used this approach (Burns & Senesac, 2005; Case, Speece, & Malloy, 2003; Speece & Case, 2001; Toregesen et al., 2001). The *Limited Norm* approach is when the distribution is limited to a particular subset of students who have received the intervention. Vellutino et al. (1996) used limited norm criteria in their study. The slopes in this study are limited to at-risk students in reading. It has also been proposed that *Benchmark* criterion can be used to determine intervention response (Fuchs, 2003). This involves meeting a criterion that is consistent with successful outcomes in the future (e.g. a fluency score that would be considered mastery level).

To evaluate response to intervention, an intervention criterion needs to be specified. One class of intervention involves more individualized attention that is different from instruction given in the general education setting. This involves *Intensive* intervention that is usually given in either one-on-one or small group settings. Both Torgesen et al. (2001) and Vellutino et al. (1996) provided the students with intense intervention services. Another class of intervention involves more general methods which can be applied to large groups of students. *General Education* interventions involve minor adaptations that can be

administered in the general education setting. Some studies have provided evidence of the effectiveness of general education interventions (Case, Speece, & Molloy, 2003; Speece & Case, 2001).

Intervention Intensity as Indicator of Response to Intervention

Concerning response to intervention, researchers have proposed that a lack of response may require increasing intervention strength by examining target variables and the intervention itself (Lentz, Allen, & Ehrhardt, 1996). The amount of intervention intensity required to produce effective outcomes for a student may provide an indicator of the need for special education services. The component of intervention intensity in the literature, however, has been described as “broad” and more research regarding this construct is needed in order to measure and select intensity variables (Barnett et al., 2004).

The basic requirements for a model implementing increasing intervention intensity are (based on Barnett et al., 2004): (1) an analysis of the intervention plan is conducted; (2) the behaviors that comprise the intervention are defined; (3) appropriate indicators of intensity are selected and a plan to measure them is developed; and (4) the extent of the episodes involving student participation and change agents are planned and checked to estimate the intervention intensity. In terms of this intensity, conclusions can be made regarding what interventions are too intense for the general education setting.

Single-case designs have used hierarchical models involving increased intensity. Rhymer, Dittmer, Skinner, and Jackson (2000) used a single-case design in order to evaluate appropriate intervention intensity needed to produce

increased fluency scores in mathematics. Other interventions with increasing intensity have created positive outcomes in other single-case designs with students exhibiting difficulties with social withdrawal (Sheridan, Kratchowill, & Elliott, 1990) and language delays (McConnell, Rush, McEvoy, Carta, Atwater, & Williams, 2002). Models implementing systematic increases in intervention need further research and support. Research regarding multiple academic and behavioral areas is also necessary to determine the models' use in educational settings.

Math Disabilities and Response to Intervention

In regards to the current research on response to intervention approaches there is a “convincing body of evidence to suggest that many children with reading difficulties can be effectively remediated by intense exposure to evidence-based reading instruction” (Gresham, VanDerHeyden, & Witt, 2005, pg. 27). However, multiple areas have remained unexplored regarding response to intervention criteria including mathematics, spelling, and writing. Research has shown that approximately 5% of the school age population experiences a disability in mathematics (Fuchs & Fuchs, 2005; Gross-Tsur, Manor, & Shalev, 1996). Although math disability is fairly prevalent there is a lack of systematic research in this area which is unfortunate because skill in mathematics is important for success in school and in the workplace (Fuchs & Fuchs, 2005). While measurement procedures needed to monitor growth have been well established, there is a need for research validating response to intervention models in this area (Vaugh & Fuchs, 2003).

Empirically Supported Math Interventions

Several intervention strategies produced have created effective outcomes for students (Fuchs, & Fuchs, 2005; Mastropieri, Scruggs, & Shiah, 1991). Reinforcement has been an effective strategy for improving mathematics calculation. Smith and Lovitt (1976) examined reinforcement's effect on arithmetic in students with learning disabilities (n=7). The results indicated that two types of reinforcement offered (extra free time or a tangible) were effective in increasing fluency rates. Luiselli and Downing (1980) used reinforcement to increase multiplication fluency in a 5th grade student with LD (as cited in Mastropieri, Scruggs, & Shiah, 1991). After the reinforcement procedures were in place, the student increased his fluency rate (the number of problems he was given to work on increased from 3 to 20).

Goal setting has also been used to increase performance in mathematics. Schunk (1985) used goal setting in students (n=30) who had been previously identified as learning disabled in mathematics. The students were then randomly assigned to one of three conditions: (1) self-set goal (students were asked to create their own, realistic goal); (2) assigned goals (the examiner would assign a particular goal to the students); or (3) no-goals. Students in both goal setting groups attempted and solved more problems than the students in the no goals group. Fuchs, Bahr, and Reith (1989) also used goal-setting as an intervention to increase mathematics performance in LD students (n=20). The researchers examined the effects of assigned versus self-selected goals. Results indicated

that the students in the self-selected goals groups were significantly more fluent than those in the assigned-goals group.

Both reinforcement and goal setting have been found to promote increased math fluency. However, more research is needed with larger samples to see if interventions that are provided in the general education classroom can promote math fluency.

Rationale

The reauthorization of IDEA now permits the use of a response to intervention (RTI) approach as an option for the identification of learning disabilities. Although the adoption of an RTI model has received some empirical support there are many issues concerning the implementation of this model that require resolution. Researchers are concerned about the legitimacy of RTI as a tool for determining learning disabilities (Gresham, VanDerHeyden, & Witt, 2005). The definition of response to intervention is not included in the IDEA regulations, nor is a procedure for its implementation described. There is a resounding need for research in this area that quantifies the concept of response to intervention and implements the approach in a succinct, scientific manner.

Although researching the literature concerning response to intervention creates an understanding as to how the approach essentially works, the available research leaves many important questions unanswered. One question involves the use of slope (rate of target skill acquisition) as a tool to differentiate among student response groups. Vellutino et al. (1996) used slopes to analyze intervention response rates in students. Although this method was found to

differentiate groups of responders, the study had a restricted sample which limited its generalizability (Fuchs, 2003). The participants in this study were all at-risk students in reading. To determine whether the students will make gains in the general education setting, it may be necessary to evaluate growth slopes from a more representative sample of students.

Using daily intervention data to produce growth slopes would allow researchers and practitioners to gain a clear understanding of each student's response to an intervention. If students are not making adequate gains with the intervention, a change is necessary to increase skill acquisition, which is discernable by examining the slope. Increasing intervention intensity can be an effective way to increase slope (and thus intervention response) (Barnett, Daly, Jones, & Lentz, 2004). However, more research is needed using general education interventions that allow for increased intensity. Intensive interventions can be time-consuming and costly (Fuchs, 2003). General education instruction, however, may not meet the needs of all students. A method that allows for increase in intensity for a general education intervention could prove to be beneficial for the majority of students. Intervention models using a hierarchy based on increasing intensity have been associated with positive outcomes in single-case designs with students exhibiting difficulties in multiple areas (McConnell et. al., 2002; Rhymer et al., 2000; Sheridan, Kratchowill, & Elliott, 1990). Research involving group design with increasing intervention intensity are needed before determining if it is an adequate predictor of intervention response and useful in determining eligibility for special education.

Current research on response to intervention approaches provides evidence to promote using a response to intervention model for reading disabilities (Gresham, VanDerHeyden, & Witt, 2005). However, multiple topics have remained unexplored regarding response to intervention criteria including mathematics, spelling, and writing. There is a need for research validating response to intervention models in these areas (Vaugh & Fuchs, 2003).

This study examines a model of response to intervention in mathematics. Within this model an effective general education intervention will be implemented to all students within one grade at an elementary school. Slope will be used to determine growth rates and learning trajectories of each student. These slopes will then be used to determine average growth in this domain and will establish groups based on the students' response to intervention. A model involving increasing intensity will be implemented to create average amounts of growth in students who have low response. This particular model is being implemented to answer whether differences exist between participant rates of learning. If there are differences, is there a pattern of normal rate of responding? Does slope or rate of learning discriminate learners? The second purpose is to determine if increasing intensity can reduce the difference between the slope of average rate learners and low rate learners.

CHAPTER II

REVIEW OF LITERATURE

Research regarding learning disabilities (LD) has been driven by social and educational necessity as a response to (1) the need to understand differences in children and adults who were identified as having specific deficits in spoken or written language while maintaining normative intellectual functioning and (2) to provide remedial services to these individuals (Lyon, 1996). Diagnostic practices for identifying LD, however, are believed to have emerged more from practice, law, and policy than from empirical support (Gresham, VanDerHeyden, & Witt, 2005; Lyon, 1996). Practices in learning disability assessment have been viewed with scrutiny due to the recent reauthorization of the Individuals With Disabilities Education Act (IDEA). The discrepancy model, in place over the last three decades, has been highly criticized. Researchers have presented several views concerning the need for change of this definition (Algozzine & Ysseldyke, 1986). A criticism surrounding discrepancy models is that they fail to differentiate low achievement from learning disabilities, overidentify students, are not implemented consistently, lack treatment validity, and delay the student's access to intervention since many approaches are based on a "wait to fail" model (Aaron, 1997; Dean & Burns, 2002; Gresham, 2001; Proctor & Prevatt, 2003). Changes in the current assessment approaches are relevant to researchers,

practitioners, teachers, and parents because these modifications may curb the overidentification of students with learning disabilities and will allow for effective interventions for those students in need of special education services.

History of Learning Disabilities Assessment

The Discrepancy Model

In 1924, J. L. Horn argued for the need of a classification system for students due to widespread failure in the school systems (Algozzine & Ysseldyke, 2001). He believed that one might have difficulties with school due to mental, behavioral, or physical reasons and proposed a categorizing system based on these groups. Around that time the discovery was made that some individuals had an average level of general overall intelligence, but had specific weaknesses in particular achievement areas; these individuals were described as having “learning disabilities” (Lyon, 1996). Although research was conducted to examine learning disabilities, this information did not have an influence on school policy until the 1960s. In 1962, Samuel A. Kirk used the term “learning disabilities” which included several groups (brain damaged, neurologically impaired, attentional disordered, etc.) with learning difficulties under one category (Oehler-Stinnett, 1986). Psychologists, parents, and educators were concerned that these children were not being effectively served by the current practices. However, learning disabled children did not qualify for special education because they did not meet the criteria for any disability. In the late 1960s and 1970s there was movement to provide services to these students and to establish a category for learning disabilities in special education (Lyon, 1996).

In 1965, Barbara Bateman suggested the concept of underachievement may be a fundamental component of the learning disability definition (Gresham, 2001). She proposed the idea of an “educationally significant discrepancy between intellectual potential and actual level of academic performance” was an indicator of LD (Gresham, 2001, pg. 10). However, this definition did not operationalize discrepancy or contribute information on how it could be measured.

The Isle of Wight studies in 1975 used the IQ-Achievement discrepancy to define learning disabilities and demonstrated how it would be used in eligibility determination (Gresham, VanDerHeyden, & Witt, 2005; Rutter, 1989). Two types of underachievement in reading were defined: general reading backwardness and specific reading retardation. Students with general reading backwardness had achievement scores that were consistent with their IQ (low achievers with no discrepancy). Students with specific reading retardation had achievement scores that were equal to or greater than two standard errors of estimate from their IQ score (low achievers with a discrepancy). In 1976, Congress commanded the Bureau of Education for the Handicapped to create eligibility criteria for specific learning disabilities. The notion of using cognitive and perceptual processing measures was rejected due to arguments that these types of assessments were psychometrically insufficient and lacking in treatment validity. However, because the eligibility criteria had to be published by December 31, 1977 (or a prevalence cap of 2% would be implemented) the discrepancy model was put into effect just prior to the deadline.

This procedure has contributed to the abundance of children who have been classified as learning disabled; the LD category accounts for 52% of the children who are served under IDEA (Gresham, 2001). Between 1976-77 and 1996-97, the number of students classified as LD had risen from 797,213 to 2,259,000, which represents a 283% increase (Gresham, 2001).

Discrepancy is still considered a fundamental component in the IDEA identification criteria for LD (Proctor & Prevatt, 2003) and the discrepancy model has been primarily used over the last two decades. Because there is no single federally mandated discrepancy model, a variety of models are used from state to state. Local Education Agencies within states may choose to use different discrepancy models. This process has been described as “confusing, unfair, and a logically inconsistent process” (Gresham, 2001, pg. 2). One of the most important issues concerning the use of the discrepancy model is that its use is not linked to producing effective interventions for students (Barnett, Daly, Jones, & Lentz, 2004). Due to this criticism, many researchers have examined a response to intervention approach.

Response to Intervention

Response to intervention (RTI) is defined as a change in performance or behavior due to intervention (Fuchs, Mock, Morgan, & Young, 2003; Gresham, 1991; Gresham, 2001). The approach does have a discrepancy base, but the discrepancy is between pre- and post-intervention levels instead of ability and achievement scores. This model deals with a *lack* of discrepancy because the goal of intervention is to produce a difference in scores. According to the

responsiveness to intervention approach, those who are nonresponsive to treatment (even after it has been extended, intensified, or changed) might have a learning disability.

The response to intervention framework was originally envisioned by Heller, Holtzman, and Messick in 1982 (Fuchs, 2003, Gresham, VanDerHeyden, & Witt, 2005). The validity of the special education eligibility criteria was examined regarding three criteria in a National Research Council Investigation: (1) the general education program's quality, (2) the ability of the special education program to create significant outcomes for its students, (3) and the precision and meaningfulness of the current process used to diagnose disabilities (Gresham, VanDerHeyden, & Witt, 2005; Vaughn & Fuchs, 2003). In regards to the current assessment process, it was determined that the assessment process must involve the student's response to instruction (Gresham, VanDerHeyden, & Witt, 2005; Vaughn & Fuchs, 2003) and that assessment must lead to decisions that could improve the student's level of academic functioning (Gresham, VanDerHeyden, & Witt, 2005). Treatment validity was regarded as a crucial component of the assessment process.

The Original Treatment Validity Model

In 1995 Fuchs operationalized the NRC criteria as a response to instruction model with four assessment phases (Vaughn & Fuchs, 2003). Phase I assesses the sufficiency of the educational environment by tracking the rate of responsiveness of all students in the classroom. Low classroom growth in comparison to other classes (in the school, district, or nation) may be indicative

of a class that needs a stronger educational program or classwide intervention (Fuchs, Fuchs, & Speece, 2002; Vaughn & Fuchs, 2003).

After it has been determined that the classroom instruction is sufficient, Phase II is used to identify students who have performance lower than their peers (Fuchs, Fuchs, & Speece, 2002; Vaughn & Fuchs, 2003). Adaptations are then put into place in the general education environment. Phase III is used to test the adaptations used to enhance the performance of this subset of children. The purpose of this phase is to determine whether the interventions in the general education environment can produce acceptable performance with this group of students. Consideration for special services is used only in circumstances where the student is not able to make adequate progress in the general education environment. The assumption within this phase is if general education adaptations “cannot produce growth for the individual, then the student has some intrinsic deficit (i.e., disability) making it difficult for him or her to derive benefit from the instructional environment that benefits the overwhelming majority of children” (Vaughn & Fuchs, 2003, pg. 138).

The model by Fuchs also includes a controversial fourth phase. In cases where the child is not making progress in the general education environment, *prior to classification*, Phase IV assessment is used to evaluate the special education program’s effectiveness for these particular students. If the special education program is found to be ineffective for the students, then “no compelling rationale exists for assigning a learning disabilities label or removing the children from the classroom for instruction” (Fuchs, Fuchs, & Speece, 2002, pg. 35). The

decision to include this phase was informed by two rationales: (1) to reduce the overrepresentation of minority students within special education and (2) to increase accountability within the special education system (Fuchs, Fuchs, & Speece, 2002). However, there were strong arguments against the addition of this phase. The first argument dealt with students who did respond to special education interventions quickly: are those students false positives (due to their quick response) or are those students the ones in need of special education services?

The second argument regarded students who would not respond to special education interventions. Perhaps those who do not respond to individualized forms of intervention in a short timeframe are the ones who require special education services the most. Should children be excluded from special education due to the fact that they did not respond to those services in a limited time frame? These would be students who exhibited difficulties within the second and third phases of assessment, indicating that they were not making adequate progress in the general education setting. Returning these students to the general education environment with no other support would not produce adequate outcomes. The difficulties with this last phase caused it to be dropped from the identification model (Vaughn & Fuchs, 2003). The treatment validity model now only contains the first three phases. Within these phases, curriculum based analysis is often used to show the growth and progress of a student. Researchers had advocated the use of curriculum-based measurement in

treatment validity models (Fuchs, Fuchs, & Speece, 2002; Vaughn & Fuchs, 2003).

Using Curriculum-Based Measurement within a Treatment Validity Model

Curriculum-Based Measurement (CBM) is a direct assessment tool used to index both academic progress and competence (Deno, Fuchs, Marston, & Shin, 2001). CBM is a set of short measures that are used to evaluate the effects of instructional programs. The most common application of CBM “requires that a student’s performance in each curriculum area be measured on a single set of global tasks repeatedly over time” (Sofie & Riccio, 2002, pg. 236).

CBM has several purposes. The first purpose involves tracking student academic ability and progress (Fuchs, Fuchs, & Speece, 2002). Deno (1985) sought to create an assessment system that would provide useful information to teachers concerning their students’ academic growth (Deno et al., 2001). CBM has been praised for its interpretability and ease (Ysseldyke, 2005). The second purpose of CBM involves answering questions concerning the ability of a program to produce academic growth. This leads to the third purpose of CBM which is improving academic programs because it provides direct information about how students are growing within the current program (Deno, 2003). CBM provides a much needed link between the student’s curriculum, intervention, and assessment.

CBM is believed to be a “hybrid” approach that includes the guidelines of Curriculum-Based Assessment while sharing some characteristics of more traditional assessments such as standardized protocol, scoring procedures,

reliability, and validity (Dombowski, 2003; Fuchs, Fuchs, & Speece, 2002). Like traditional measurement, CBM evaluates an extensive range of skills weekly, each measurement determined by which skills were covered that week. Weeks of repeated measurement provide alternate forms of measurement for single skills, such as addition accuracy, oral reading fluency, etc. (Fuchs, Fuchs, & Speece, 2002). It also provides data concerning maintenance of particular skills that are no longer being directly addressed in the curriculum.

CBM scores can be viewed as performance indicators because they can create ranges of scores across students of the same age and allows for rank ordering of these individuals (Deno et al., 2001). The CBM score is representative of the student's global level of proficiency in a particular academic area. This in turn can be used to compare a student against his/her peers and can indicate deficiencies in particular academic domains.

The treatment validity model relies on using CBM as an assessment tool for several reasons. The first is that it can operationalize the three assessment phases proposed by Fuchs by documenting growth between students (Vaughn & Fuchs, 2003). It has the ability to measure the classroom's level of instruction (Phase I), identify those students whose level of growth is below that of their peers (Phase II), and document the rate of response to classroom adaptations (Phase III). CBM can demonstrate a student's growth and can inform researchers, practitioners, and teachers about the effectiveness of a particular intervention (Marston, Fuchs, & Deno, 1986; Fuchs, Fuchs, & Hamlett, 1989). This information can aid in decisions about interventions—whether the current

intervention needs to be intensified, faded, or replaced depending on the student's current level of performance.

CBM also has the ability to distinguish inadequate instruction versus inadequate individual learning (Fuchs, Fuchs, & Speece, 2002). If CBM is administered to all the students in a particular grade, one could determine differences in progress between classrooms. For example, if most of the slopes in Classroom A are generally low in comparison to Classroom B, this is possibly indicative of poor instruction in Classroom A. However, if a student from Classroom B (which has been found to have strong instruction) has a slope much lower than the other students, this may indicate a learning problem.

Although researchers have advocated the effectiveness of CBM's progress monitoring and assessment strategies for students with disabilities (Fuchs, Roberts, Fuchs, & Bowers, 1996; Vaughn & Fuchs, 2003) there is a lack of systematic application of these tools as a means of identifying students with learning disabilities.

Speece and Case (2001) provide an excellent example of curriculum-based measurement's use in identifying children with learning disabilities. In this study, the researchers use a dual-discrepancy model. This model uses curriculum-based measures (CBM) to identify children whose performance is below that of their classmates (Fuchs, Fuchs, & Speece, 2002; Speece, Case, & Molloy, 2003). In this model CBM creates a slope in which to judge the child's responsiveness to the general education environment. The discrepancy between the child's level and slope indicates if there is a need for specialized instruction.

For a child to be placed into special education services, they must show a dual-discrepancy, meaning that the child performs below the level of the classroom and shows a learning rate below same-age peers (Gresham, 2001).

Speece and Case (2001) used Letter Sounds Fluency and Oral Reading Fluency probes to ascertain students' reading levels. Children in the lowest 25% of their classroom were deemed at risk for reading failure. A contrast group (called the "purposive sample") included 5 students from each classroom (2 students from the median and 1 each from the 30th, 75th, and 90th percentiles). Each student was assessed using the reading probes, four subtests from the Wechsler Intelligence Scale for Children—Revised, two subtests from the WJ-R, a phonological processing battery, and a Rapid Automatized Naming task. The students in the at-risk group were then categorized into one of three groups: CBM dual discrepancy (CBM-DD), regression based IQ-reading achievement discrepancy (IQ-DS), and low achievement (LA). The students in the CBM-DD group (n=47) were based on scores of 10 CBM Oral Reading Fluency Probes administered over the school year. These students had a slope across the year and level of performance at the end of the year that was more than 1 standard deviation below the slope and level of their peers. Students in the IQ-DS group (n=17) were those who had reading achievement scores 1.5 or more standard errors of predication below their predicted achievement. Students in the LA group (n=28) were those who had reading achievement score of less than 90.

Analyses showed that the children who were identified as CBM-DD were younger and had more difficulties on phonological processing tasks and teacher

ratings of academic competence and social behaviors than those in the IQ-DS and LA groups. It was found, however, that single-point measures of reading fluency did not accurately identify students with reading difficulty. The readers suggested that repeated evaluation may be necessary for valid eligibility determination. Although the CBM-DD approach may seem more cumbersome than traditional IQ-Achievement discrepancy model, it “may be the price of valid procedures” (Speece & Case, 2001, pg. 747). However, before adopting a response to intervention approach using CBM, further research of methods and procedures is needed.

Response to Intervention Models

In its most basic form, RTI can be described in five steps (Fuchs et al., 2003). First, students are given effective classroom instruction in the general education setting. Second, the progress in the general education setting is monitored. Third, those students who are not making progress in the general education receive interventions from their teacher or another individual. Fourth, the students’ progress with the intervention is monitored. Fifth, those students who are not responding to the intervention either qualify for special education services or an evaluation (which may later lead to special education services). Although the RTI model seems very simplistic, there are several ways to implement this model. Currently, there are two fundamental frameworks for RTI models: (1) the problem solving approach and (2) the standard protocol approach (Fuchs et al., 2003).

Problem Solving

Problem solving approaches have been beneficial to consultation and assessment practices in educational environments (Telzrow, McNamara, & Hollinger, 2000). The problem solving approach stems from the behavioral consultation literature which was first defined by Bergan (1977) and Bergan and Kratochwill (1990) (Fuchs et al., 2003; Gresham, VanDerHeyden, & Witt, 2005). The most prominent feature of the problem solving model is that it is inductive, meaning that “no student characteristic dictates a priori what intervention will work” (Fuchs et al., 2003, pg. 160). Instead of assuming a particular intervention will be beneficial for a student the model uses a four step process involving problem identification, problem analysis, plan implementation, and problem evaluation to guide the intervention process (Telzrow, McNamara, & Hollinger, 2000). The purpose of the problem identification phase is to define an observable problem behavior and measure its frequency, duration, or intensity (Fuchs, Mock, Morgan & Young, 2003). In the problem analysis phase, the consultant identifies variables that are contributing to the problem and then develops a plan to remediate the current behavior. During the plan implementation phase the consultant monitors the intervention and its treatment integrity to make certain that the intervention is run accurately and consistently. In the final phase, problem evaluation, the consultant meets with the teacher to evaluate the intervention’s effectiveness. If the intervention is not effective, it is either modified or replaced. Throughout all phases data are collected to evaluate the student’s response to the intervention. The result is an intervention that is specifically tailored to the student’s needs.

One of the main purposes of the problem solving model is to identify problems in “context of person-environment situations rather than attributing them to putatively fixed student characteristics” (Fuchs et al., 2003, pg. 160). It involves collaboration between the consultant, teacher, and student. The consultant’s effect on the student is indirect, meaning that the consultant works through the teacher to produce outcomes for the student, usually by means of an intervention. The problem solving approach is very popular among practitioners, and variants of this model have been used in the form of pre-referral intervention teams in several areas across the country (Ohio’s Intervention Based Assessment, Pennsylvania’s Instructional Support Teams) (Fuchs, Mock, Morgan & Young, 2003). Some school districts are also using the approach in eligibility decision making (Heartland Educational Agency in Iowa, and Minneapolis Public Schools).

Although the use of problem solving approach is used among practitioners in some districts, there are still concerns regarding its outcomes. There is inadequate evidence as to the effectiveness of these approaches in Ohio, Pennsylvania, Heartland, and Minneapolis and little research to advocate the usefulness of the model (Fuchs et al., 2003). There are some concerns as to whether the interventions within these models are implemented with integrity and are producing effective outcomes for its students. The behavioral consultation literature has provided evidence that teachers can effectively implement interventions with integrity. Researchers have examined this effect of performance feedback and have found it to result in high stable intervention

implementation for the majority of teachers (Mortenson & Witt, 1998; Noell, Duhon, Gatti, & Connell, 2002).

Even though the literature demonstrates that teachers can provide interventions with integrity, the research also indicates that certain structures need to be in place to increase accountability. However, “practitioners of the problem-solving approach typically have not produced fidelity of implementation information or they have documented low levels of implementation accuracy” (Fuchs et al., 2003, pg. 167). There is a need for more research regarding these RTI models that are currently in place to demonstrate that they can provide timely and effective interventions.

Standard Protocol Approach

The second RTI approach uses the same empirically validated intervention across all students experiencing difficulties in a particular academic domain (e.g., addition) (Fuchs et al., 2003, Gresham, VanDerHeyden, & Witt, 2005). Therefore large groups of students would be able to benefit from an effective treatment protocol (Fuchs et al., 2003). One of the main criticisms of the IQ-Achievement discrepancy model is that it fails to differentiate between struggling students who have had inadequate instruction or exposure to educational materials versus those who have “true” learning disabilities. This approach would be helpful in differentiating between those students who may have problems due to actual deficits than lack of practice or exposure to instruction (Vellutino et al., 1996). The assumption of the standard protocol approach is if a student responds to an intensive treatment trial then he/she does

not have a disability, the intervention remediates the problem, and the student can return to the general education classroom (Fuchs, Fuchs, & Compton, 2004).

Vellutino et al.'s (1996) study is regarded as an excellent example of the implementation of a standard protocol model (Fuchs et al., 2003, Gresham, VanDerHeyden, & Witt, 2005). The researchers asked first-grade teachers to rate each of their students in regards to reading ability at the beginning of the school year excluding those who had severe hearing or vision problems, frequent ear infections, or severe emotional problems. Students who took daily medication, spoke English as a second language, had limited intellectual ability, and had a diagnosis of a pervasive neurological disorder were also excluded from the study. The researchers then randomly selected students who were identified as normal readers. The students in these groups were then administered the Word Attack and Word Identification subtests of the Woodcock Reading Mastery Tests—Revised (WRMT-R; Woodcock, 1987). Students scoring at or below the 15th percentile composed the poor readers group (n=118). Those who scored at or above the 40th percentile composed the normal readers group (n=65).

The poor readers were divided into two groups: the tutored group and the nontutored group. Those in the tutored group (n=76) were given daily one-on-one tutoring for 30 minutes per day for 15 weeks. Those who did not receive tutoring (n=42) received remediation through their school, but interventions varied from one-on-one instruction to small group instruction. The students who received small group instruction served as a contrast group to examine the effectiveness

of the intervention. These normal readers group received no intervention except for the reading instruction supplied by their classroom teacher.

The students were repeatedly administered the WRMT-R tests throughout the study. These assessments were used to calculate growth curves for each student. Slopes derived from linear regression were rank ordered and used to place students in responsiveness groups. The data indicated four levels of responsiveness: “very limited growth”, “limited growth”, “good growth”, and “very good growth”. Two-thirds of the tutored group established either “good growth” or “very good growth.” From these results, Vellutino et al. proposed that these students did not have reading deficits but were impaired by instructional deficits. The students who had made “limited growth” and “very limited growth” were described as “difficult to remediate.” However, the researchers reported these results were inconclusive due to the fact that no comparison was made between the cognitive abilities of the children who were readily remediated versus those who were difficult to remediate. The study does, however, illustrate the use of a standard protocol approach.

Torgesen et al. (2001) also used a standard protocol in order to provide remedial instruction to students with severe reading disabilities. Participants (n=50) were students from LD classes. Prior to the intervention implementation, the children were pre-tested with a battery which included: measures of phonological processes, eight measures from the WRMT-R, measures of other academic skills (spelling and calculation subtests), measures of expressive and

receptive language, an IQ test, teacher behavior checklists, and an assessment of fine-motor function.

Children in the sample were randomly assigned to one of two groups: the Auditory Discrimination in Depth Program (ADD) and the Embedded Phonics Program (EP). The ADD program focuses on (1) the ability of the student to discriminate phonemes (by teaching kinesthetic, auditory, and visual features associated with phonemes); (2) the aptitude of the student to monitor and represent sounds in spoken syllables; and (3) the capability of the student to monitor these skills and use self-correction (Lindamood & Lindamood, 1998). The EP program gives the student direct instruction in word-level reading skills and provides them opportunities to read and write meaningful text. The students were given daily one-on-one tutoring two 50 minute sessions per day for 8 to 9 weeks (for a total of 67.5 hours of instruction). Two-three weeks after the intensive intervention phase, the children were given the same phonological awareness measures, measures of other academic skills, and expressive and receptive language measures.

After the intense intervention, each student was provided with generalization training for 8 weeks. The teacher who worked with the child during the intensive phase going into the LD class for one 50 minute session each week worked with the child using classroom materials. Follow-up measures were then administered again at 1 and 2 year intervals following the original post-test in order to monitor growth in both reading and language abilities.

Both interventions proved effective in improving generalized reading skills over the 2 year follow-up. Growth on the WRMT-R was statistically significant at the 2-year follow up, $F(2, 94) = 8.6, p < .01$. Although children's average scores on the measures of reading accuracy and comprehension were in the average range at the end of the follow-up, measures of reading rate still showed severe impairment for most children. However, 1 year after the intervention, 40% of the children were found to no longer need special education services.

Vaugh, Linan-Thompson, and Hickman (2003) also used a standard protocol approach as a means of identifying students with reading disabilities. Second grade students ($n=45$) were identified as at risk for a reading disability. The researchers used a two-tiered process for identification purposes. The students were first nominated by teachers for participation if they were reading below grade level. Then the students were assessed using the screening portion of the Texas Primary Reading Inventory (TPRI); those who met the at-risk criteria were included in the study. Prior to intervention the students were administered a WRMT-R (Word Attack and Passage Comprehension subtests), a Comprehensive Test of Phonological Processing (CTOPP), and the Test of Oral Reading Fluency (TORF). This battery was given on four occasions: prior to intervention and after each of three 10-week intervals. During these intervals, the students who met exit criteria were dismissed from the intervention, but still participated in the TORF assessment.

The intervention targeted five elements of reading development: phonemic awareness, phonics and mastery of sound-letter relationship and word families,

reading fluency, comprehension, and spelling. Students received 35 minutes of small group instruction daily. Reading fluency was monitored weekly by assessing each student's words read correctly per minute.

At each intervention interval, students who met criteria were discontinued from further intervention. Those who exited at the first intervals were called the early exit group (n=10). Those who exited at the second interval were referred to as the midterm exit group (n=14). Those who exited at the final interval were the late exit group (n=10). The remaining students who did not meet exit criteria after the intervention period (n=11) represented less than 25% of the students determined at risk for a reading disability. These students would be considered for special education. Measures that predicted membership in this no exit group were pretest scores on fluency, passage comprehension, and rapid naming tasks.

Each of these standard protocol approaches listed above transform an identification process into a process of prevention. Not only did they provide more stringent criteria for identification, none of the students in the study had to wait to receive an intervention service. This is a benefit of the response to intervention model.

Comparison of Problem Solving and Standard Protocol Approaches

Although the overall goal of both the problem solving and standard protocol approaches is the same, the reasoning and structure behind the models are different. The problem solving approach uses inductive reasoning, meaning that arguments are made through observation of the student's behavior. The

decision-making process is based on the data that are collected throughout the problem solving process. Therefore, each intervention is unique in that it is tailored for the student. Reasoning in the standard protocol approach is deductive, meaning that arguments in this approach are made based on empirically-validated research. The decision-making process in this approach is based on interventions that worked in the past for large numbers of students. Therefore, the same intervention is used for all students with similar concerns.

There are advantages and disadvantages for each model. The problem solving approach is very sensitive to individual differences. This in turn could produce very effective outcomes for those students. However, because this process involves working with individual students it is more difficult to implement. Procedures requiring more time and effort may not be readily accepted. Ysseldyke (2001) states “Change is difficult and more political than data-based...and while change is difficult, change requiring extra work is next to impossible” (pg. 300). Though this model may be able to produce beneficial outcomes for students, more research streamlining the approach and documenting its effectiveness is required. On the other hand, the standard protocol approach has greater quality control. This approach could produce outcomes for large groups of students across varying levels of performance. However, in this model students do not receive individualized instruction. The students most in need of intervention may not be able to receive the intensity they require to make progress. Both models have yet to produce research

concerning how these approaches will be adopted on a large scale (Fuchs et al., 2003).

Intervention Response Options in RTI

There are several approaches of RTI that have been researched in the last decade. These methods view differences involving timing of measurement (final status, growth, or dual-discrepancy), standards for designating responsiveness, and the nature of intervention (intensive versus general education) (Fuchs, 2003). Each of these options has its own implications and issues.

Measurement Timing

There have been three types of timing procedures documented in the RTI literature: final status, growth, and dual discrepancy. Final status involves testing the student at the end of the intervention period, with the student's responses based on their post-intervention outcome (Fuchs, 2003). Growth models measure the student's progress periodically throughout the intervention in order to ascertain how much learning took place. Dual discrepancy models determine if the student's level and slope is significantly lower than that of his/her peers.

Final status models. Final status is considered a "straightforward" method in determining whether the student has made adequate response to the intervention (Gresham, VanDeyHeyden, & Witt, 2005). The main component of this procedure is determining the criterion for sufficient response. Those who do not meet that criterion would be regarded as needing special education services. No strict guidelines for final status have been determined. This criteria could be

based on normative scores, such as meeting the 25th percentile or greater on the assessments (Fuchs, 2003).

Torgesen et al. (2001) used a final status approach in determining which students had responded adequately to their intervention. Students who failed to achieve WRMT-R and GORT-III scores of 90 or greater were defined as having inadequate response. Forty percent of their sample (all of whom were identified as learning disabled) met this criteria and were identified as no longer needing special education services.

Vaughn, Linan-Thompson, and Hickman (2003) also used a final status approach. Although progress monitoring data were collected throughout the study, students needed to meet exit criteria to be determined responsive to the intervention. This exit criteria were a passing score on the TPRI, median score or higher performance on the TORF, and a fluency score of 50 words correct per minute (with fewer than 5 errors). Using these criteria, over 75% of the students were regarded as adequately responding to the intervention.

Although both of these studies show succinct and clear methods for determining responsiveness, they fail to look at student growth. Some of the students in the study may have started at a lower level and may have progressed more than those who had higher pre-intervention scores. Growth may be a better indicator of intervention response.

Growth models. Growth models seek to produce a discernable difference between pre-intervention and post-intervention levels of performance (Gresham, VanDerHeyden, & Witt, 2005). This model provides a way to view the student's

current trajectory in the target area. Vellutino et al. (1996) used growth to analyze intervention response rates in students by repeated administration of the WRMT-R tests. These assessments were used to calculate growth curves for each student. Slopes derived from linear regression were rank ordered and used to place students in responsiveness groups. The data indicated four levels of responsiveness: “very limited growth”, “limited growth”, “good growth”, and “very good growth”. Two-thirds of the tutored group established either “good growth” or “very good growth” and were regarded as those who responded adequately from treatment.

Although this method does an excellent job of tracking each student’s progress, its reliance on a limited sample to determine response is problematic (Fuchs, 2003). The sample in this study was comprised of students who were at risk in reading. A sample that better reflects the population of general education needs to be studied.

Dual discrepancy model. To address some of the issues generated by the above models, Fuchs and Fuchs (1998) proposed a dual discrepancy model which relies on evaluation of a student’s level and slope in contrast to their peers. Case, Speece, and Molloy (2003) tested a response to intervention model using general education interventions and the dual-discrepancy approach. They classified first and second graders into three responsiveness groups—Frequently Dually Discrepant (FDD), Infrequently Dually Discrepant, (IDD), and Never Dually Discrepant (NDD)—using CBM measures of Letter Sounds Fluency and Oral Reading Fluency. Children were identified as “at-risk” if these probes placed

them in the lowest 25%; the at-risk group consisted of 36 children who remained in the study for 3 years. The discrepancy groups were compared against controls (peers who scored on the 30th, 75th and 90th percentiles) on measures of phonological processing, behavior, and instructional context. The students were tested repeatedly for a duration of three years.

Repeated measures ANOVA yielded significant main effects for discrepancy group on measures of phonological processing and behavior; in cases involving a significant group effect, the FDD group performed significantly worse than the NDD group. There was no significant interaction of group by instructional context. Overall, the FDD group demonstrated greater learning problems in the general education setting. The researchers concluded that a response to instruction model can be implemented within a general education setting to identify students who are having difficulties and require more instruction.

Burns and Senesac (2005) also examined a dual discrepancy approach in assessing response to reading intervention. Participants (n=146) were students who were experiencing reading difficulties and who scored below the 25th percentile on a group test of reading. The students were provided one of two interventions in this study: the Help One Student to Succeed program (HOSTS) or Title I intervention services. Student response to the intervention was measured using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency (ORF) probes. Students were placed into DD groups based on this measure. The GORT-4 was used to determine level of reading

proficiency between the groups. These measures were administered twice (both pre- and post-intervention). These scores were then used to place students in to dual discrepancy groups based on the normative criteria by using the 25th, 33rd, 50th percentile ranks, and less than one standard deviation below the mean. Students whose post-intervention DIBELS score fell within the at-risk criterion (less than 20 words correct per minute) *and* whose fluency change score who fell below the non-responsiveness criterion (the above percentile ranks) were classified as dually discrepant.

Once the groups were established, the GORT-4 scores were analyzed to compare student groups. Results indicated that three of the four models adequately differentiated between those who were DD and non-DD (the one standard deviation group did not differentiate between the groups).

A major limitation of this study is that only two data points were used for criteria. It did not use slope data from weekly probes (which may have provided a better indicator of growth). Decisions about a student's eligibility for special education services should not be made on the basis of one or two point-in-time assessments of student characteristics (Barnett et al., 2004). Perhaps using slope of learning is a better indicator of overall performance.

Standards for Designating Responsiveness

Despite the measurement used a second standard must be applied to the measurement to create a cutpoint for differentiating responders from nonresponders (Fuchs, 2003). This criterion can be gauged from viewing the full distribution of student functioning, which is referred to as the *Normative*

approach. An example of this is using the 25th percentile relative to a sample which has a full range of student performances included. Several studies have used this approach (Burns & Senesac, 2005; Case, Speece, & Malloy, 2003; Speece & Case, 2001; Torgesen et al., 2001).

When this distribution is limited to a particular subset students who receive the intervention, this is referred to as *Limited Norm*. Vellutino et al. (1996) used limited norm criteria in their study. The slopes in this study are limited to at-risk students in reading.

It has also been proposed that *Benchmark* criterion can be used to determine intervention response (Fuchs, 2003). This involves meeting a criterion that is consistent with successful outcomes in the future. An example of this would be reading fluency at the mastery level by the end of first grade. However, there is a lack of research using benchmark criterion to gauge responsiveness.

Nature of Intervention

To evaluate response to intervention, an intervention criterion needs to be specified. One class of intervention involves more individualized instruction than the curriculum administered in the general education setting. This involves *Intensive* intervention given in either one-on-one or small group settings. Both Torgesen et al. (2001) and Vellutino et al. (1996) provided students with intense intervention services.

Another class of intervention involves more general methods which can be applied to large groups of students. *General Education* interventions involve minor adaptations administered in the general education setting. Intensive

remediation can be time consuming and assumes that progress provides evidence that the student's difficulties were caused by poor instruction instead of an actual reading deficit (Fuchs, 2003); although some students may progress through intensive intervention, they may still have difficulties in a general education setting (Fuchs, 2003). To address these issues a remediation should be grounded in general education where interventions used to measure responsiveness have proved to be efficacious for a majority of students (Fuchs, 2003). Some studies have provided evidence of the effectiveness of general education interventions (Case, Speece, & Molloy, 2003; Speece & Case, 2001).

Perhaps a model employing both general education interventions and intensive interventions (for those who are deemed more at-risk) would be beneficial). This would start with the application of a general education intervention for all students and the same intervention would be applied with increasing intensity to students most at-risk for learning difficulties.

Intervention Intensity as Indicator of Response to Intervention

The component of intervention intensity in the literature has been described as "broad" and more research regarding this construct is needed to measure and select intensity variables (Barnett et al., 2004). Concerning response to intervention, it has been proposed that a lack of response may require increasing intervention strength by examining target variables and the intervention itself (Lentz, Allen, & Ehrhardt, 1996).

When measuring intervention intensity variables in context there are two components that need assessment. The first involves "socially valid child out-

come variables that can be measured repeatedly over time” (Barnett et al., 2004, pg. 68). The use of curriculum-based measurement can be used to gauge important academic outcomes. CBM is also useful in determining progress over time.

The second component involves selecting variables that allow for the intensity of the intervention to be easily quantified. The basic requirements for a model implementing increasing intervention intensity are (based on Barnett et al., 2004): (1) an analysis of the intervention plan is conducted; (2) the behaviors that comprise the intervention are defined; (3) appropriate indicators of intensity are selected and a plan to measure them is developed; and (4) the extent of the episodes involving student participation and change agents are planned and checked to estimate the intervention intensity. In terms of this intensity, conclusions can be made regarding what interventions are too intense for the general education setting.

Rhymer, Dittmer, Skinner, and Jackson (2000) used a single-case design to evaluate appropriate intervention intensity needed to produce adequate outcomes. The researchers used an increasing intervention intensity design involving combined timings, peer tutoring, positive-practice overcorrection and performance feedback to increase mathematics fluency. The outcomes from a brief, fluency intervention were minor for the students (n=4). However, considerable improvement was made by 3 of the students when a performance feedback component was added to the fluency intervention.

Other interventions with increasing intensity have created positive outcomes in single-case designs with students exhibiting difficulties with social withdrawal (Sheridan, Kratchowill, & Elliott, 1990) and language delays (McConnell, Rush, McEvoy, Carta, Atwater, & Williams, 2002). Although these studies validate models using increasing intensity they have only been conducted using single-case design. Research involving group designs with increasing intervention intensity are needed before determining if it is an adequate predictor of intervention response and useful in determining eligibility for special education. Research regarding multiple academic and behavioral areas is also necessary to determine the models' broad capacity for application in educational settings.

Response Intervention Models By Academic Area

In regards to the current research on response to intervention approaches there is a "convincing body of evidence to suggest that many children with reading difficulties can be effectively remediated by intense exposure to evidence-based reading instruction" (Gresham, VanDerHeyden, & Witt, 2005, pg. 27). However, multiple areas have remained unresearched regarding response to intervention criteria including mathematics, spelling, and writing.

Approximately 5% of the school age population experiences a disability in mathematics (Fuchs & Fuchs, 2005; Gross-Tsur, Manor, & Shalev, 1996). Although math disability is fairly prevalent there is a lack of systematic research in this area which is unfortunate because skill in mathematics is important for success in school and in the workplace (Fuchs & Fuchs, 2005). Although measurements procedures used to monitor growth have been well established,

there is a need for research validating response to intervention models in this area (Vaugh & Fuchs, 2003).

Math Interventions

To evaluate the response to mathematics intervention, it is important to present research of interventions that have proven effective in this area. Several math interventions have been found to create effective outcomes in students (Mastropieri, Scruggs, & Shiah, 1991).

Reinforcement

Reinforcement has been an effective strategy in improving mathematics calculation. Smith and Lovitt (1976) examined the effect of reinforcement on arithmetic fluency in students with learning disabilities (n=7). Two types of reinforcement were used to increase fluency: contingent free time (student could earn free time when criteria was met) and contingent toy (student could earn a toy when criteria was met). The students earned points based on their math fluency. A ratio was calculated for the students that would determine how many problems they would have to finish to earn a toy or free time. (e.g. every two problems calculated correctly would earn one minute of free time). Both reinforcement strategies were effective in increasing fluency rates.

Luiselli and Downing (1980) used reinforcement to increase multiplication fluency in a 5th grade student with LD (as cited in Mastropieri, Scruggs, & Shiah, 1991). The student received praise for correctly completing the problems for the day. With the reinforcement procedures in place, the student increased his fluency rate from 3 to 20 problems. Although reinforcement was found to be

effective in these single-case designs, more research is needed to establish its impact on mathematics fluency in group designs.

Goal Setting

Schunk (1985) used goal setting in students (n=30) who had been previously identified as learning disabled in mathematics. Each student received training sessions involving subtraction skills for 45 minutes per day for 5 consecutive school days. The students were then randomly assigned to one of three conditions: (1) self-set goal (students were asked to create their own, realistic goal); (2) assigned goals (the examiner would assign a particular goal to the students); or (3) no-goals. Analyses revealed that the self-set goals group performed higher than the assigned goals and no goals groups. Students in both goal setting groups attempted and solved more problems than the students in the no goals group.

Fuchs, Bahr, and Reith (1989) also used goal-setting as an intervention to increase mathematics performance in LD students (n=20). The researchers examined the effects of assigned versus self-selected goals and contingent versus noncontingent gameplay conditions. The students were assigned randomly to one of four groups: (1) assigned goal/noncontingent gameplay; (2) self-selected goal/noncontingent gameplay; (3) assigned goal/contingent gameplay; and (4) self-selected goal/contingent gameplay. Computer assisted drill and practice was provided to all students. The students in the self-selected goals groups were significantly more fluent than those in the assigned-goals

group. There were no significant effects of the contingent play versus noncontingent play groups.

Both reinforcement and goal setting were found to promote increased math fluency in each of the studies presented. However, more research is needed with larger populations to see if interventions that are provided in the general education classroom can promote math fluency.

Rationale

The reauthorization of IDEA now permits the use of a response to intervention (RTI) approach as an option for the identification of learning disabilities. Although the adoption of an RTI model has received some empirical support (Case, Speece, & Molloy, 2003; Gresham, 2001; Gresham, VanDerHeyden, & Witt, 2005; Speece, Case, & Molloy, 2003; Torgesen et al., 2001; Vellutino et al., 1996) there are many issues concerning the implementation of this model that require resolution. Researchers are concerned about the legitimacy of RTI as a tool for determining learning disabilities (Gresham, VanDerHeyden, & Witt, 2005). The definition of response to intervention is not included in the IDEA regulations, nor is a procedure for its implementation described. There is a resounding need for research in this area that quantifies the concept of response to intervention and implements the approach in a succinct, systematic and defensible manner.

Although researching the literature concerning response to intervention creates an understanding as to how the approach essentially works, there are still many important questions unanswered. One concern involves the use of

slope (of intervention growth) as a tool to differentiate among student response groups. Vellutino et al. (1996) used slopes to analyze intervention response rates in students. Although this method was found to differentiate groups of responders, its reliance on a restricted sample limited the study's generalizability (Fuchs, 2003). The participants were limited to at-risk students in reading. To determine whether the students will make gains in the general education setting, it may be necessary to evaluate growth slopes from a more representative sample of general education students.

More research is needed using general education interventions that allow for increased intensity. Intensive interventions can be time-consuming and costly (Fuchs, 2003). General education interventions, however, may not meet the needs of all students. A method that allows for increase in intensity for a general education intervention could prove to be beneficial for the majority of students. Intervention models using a hierarchy based on increasing intensity have been associated with positive outcomes in single-case designs with students exhibiting difficulties in multiple areas (McConnell et. al., 2002; Rhymer et al., 2000; Sheridan, Kratchowill, & Elliott, 1990). Research using group designs with increasing intervention intensity are needed before determining if it is an adequate predictor of intervention response and useful in determining eligibility for special education.

There is a much evidence to promote using a response to intervention model for reading disabilities (Gresham, VanDerHeyden, & Witt, 2005). However, multiple topics have remained unexplored regarding response to

intervention criteria including mathematics, spelling, and writing. It has been shown that approximately 5% of the school age population experiences a disability in mathematics (Fuchs & Fuchs, 2005; Gross-Tsur, Manor, & Shalev, 1996). Although measurements procedures to monitor growth have been well established, there is a need for research validating response to intervention models in this area (Vaugh & Fuchs, 2003).

This study will examine a model of response to intervention in mathematics. Within this model an effective general education intervention will be implemented to all students within one grade at an elementary school. Slope will be used to determine growth rates and learning trajectories of each student. These slopes will then be used to determine average growth in this domain and will establish groups based on the students' response to intervention. A model involving increasing intensity will be implemented to create average amounts of growth in students who have indicated low response. This particular model is being implemented to answer crucial questions regarding response to intervention.

Research Questions and Hypotheses

One primary question to be answered by this study is whether differences exist between participant rates of learning. If there are differences, is there a pattern of normal rate of responding? Can slope be used to determine learning outcomes? The second question regards intervention intensity. Can increasing intervention intensity reduce the difference between those who have an average learning slope and those who learn at a slower rate (those who are more

resistant to intervention)? The following list summarizes the questions that will serve to answer those primary concerns:

Research Question 1: Does rate of learning vary in a normative sample of students?

It is hypothesized that fluency data will indicate a distribution approximating a normal curve with low, average, and high rate learners. The null hypothesis states that there will be no significant difference in learning rates.

Research Question 2: Can increasing intensity reduce the difference between the slope of average rate learners and low rate learners?

It is hypothesized that increasing intervention intensity will reduce the difference in slope between average rate learners and low rate learners. The null hypothesis states that increasing intervention intensity will not reduce the difference in slope between average rate learners and low rate learners.

CHAPTER III

METHODS

Participants and Setting

The participants in this study were 5 teachers and 71 general education students from an elementary school in a Southwest rural community. The elementary school serves approximately 600 students in kindergarten through fourth grade. Superintendent, principal, and teacher consent were obtained prior to data collection procedures (See Appendix A for Research Prospectus). The first intervention phase (Establishing Rate of Response) included all of the students from the 5 second grade classrooms. The second intervention phase (Rate of Response Matching) included 4 students: Colton, Mallory, Kenneth, and Zeke (pseudonyms) who exhibited low response rates during the first intervention phase. Parent consent and child assent were obtained for the 4 students included in phase two intervention procedures. Both parent permission and child assent forms stated that the student could withdraw permission at any time to remove themselves from the research project.

Phase one intervention procedures were conducted by the experimenter and research team members in the classroom setting during the scheduled mathematics instruction. Phase two intervention procedures were conducted by the experimenter and team members in the Title 1 reading room, a small classroom with a single table and chairs.

Materials

Single-Skill Probe

Materials for this study consisted of mathematics probes comprised of single digit subtraction problems. The probes were 1-digit by 1-digit subtraction problems, subtraction from 9, presented vertically in eight rows and eight columns with sixty-four problems per page. This skill has been identified as a second grade Priority Academic Student Skill (Oklahoma PASS Standards), which are the state curriculum standards for the district in which the study was conducted. The subtraction probes were generated using a Microsoft Excel worksheet (See Appendix B). The RANDBETWEEN function in Excel was used to generate random numbers between 0 and 9 for subtraction problems so randomized probes could be created for each session of the study.

Reinforcements

Reinforcements were used during the intervention phases of the study in an attempt to ensure high levels of effort throughout. The teachers were asked to identify acceptable reinforcers for the classroom. The reinforcement box included items such as pencils, stickers, candy, and bookmarks.

Dependent Variable

The dependent variables in this study were multiplication fluency and responsiveness. Fluency is defined as the number of total digits completed accurately during the two minute assessment (Shinn, 1989). Fluency will be assessed using single digit multiplication probes. A digit is deemed correct if it is located in the correct column of the answer (Skinner, Turco, Beatty, & Rasavage, 1989). For example, an answer of “45” to the question “5 x 9” would receive 2 points since both digits are in the correct column. An answer of “43” would receive 1 point since 1 digit is in the appropriate column. An answer of “33” would receive no points since both digits are incorrect.

Responsiveness was examined to determine if slope could discriminate between learners and if increasing intervention intensity could reduce the difference between the slope of average rate learners and low rate learners. Responsiveness was operationalized as the slope of fluency scores. The fluency scores from each intervention point were graphed (using a line graph in Excel) to create a slope for each student. The SLOPE worksheet function was then used in Excel to calculate the slope of the linear regression line through each student’s data points. In Excel, the SLOPE function calculates the vertical distance divided by the horizontal distance between any two points on the line, thus calculating the rate of change along the regression line. Calculating the slope of each student’s fluency serves two purposes: (1) assessing the full range of students allowed the experimenter to specify a normative profile of growth (or average rate

of response) and (2) once that normative rate of response was determined, other students' growth was compared to that normative profile.

Fluency Intervention

The purpose of the intervention was to improve fluency of subtraction from 9, a keystone skill that is an essential component of the second grade curriculum. The fluency intervention incorporated both goal setting and reinforced practice, two methods that have shown to be effective in maintaining accuracy rates (Freeland & Noell, 1999) and promoting fluency (Mastropieri, Scruggs, & Shiah, 1991) in elementary mathematics. This intervention involved rewarding the student for exceeding a fluency goal based upon previous performance. For example, a student would be able to take a reward from the reinforcement box if he/she exceeded the previous day's performance by at least 1 point (in this case, 1 more digit correct).

Overview of Procedures and Experimental Design

Prior to the experiment, approval was obtained from the institutional review board (IRB) of Oklahoma State University (OSU) (See Appendix). This study involved four data collection phases. The first, *Pre-Intervention Phase*, involved screening procedures to determine the sample of students used for the study and baseline procedures. The second, titled *Intervention Phase One*, involved the use of a fluency intervention across the entire sample to establish a

rate of intervention response. This phase used a non-experimental design. The third phase, *Intervention Phase Two*, entailed using the same intervention at increasing intensity levels to improve the response of students who responded poorly to the initial intervention phase. During this phase, student response to increasing intervention intensity was evaluated within a multiple baseline design across subjects. The fourth, *Post-Intervention Phase*, involved follow-up data collection procedures designed to evaluate maintenance of fluency rates.

Pre-Intervention Phase: Establishing Performance Levels

Screening. School-wide screening procedures were used to identify a subset of students in which a single fluency intervention was deemed appropriate. Deficiencies in calculation skills were identified across the multiple grades screened. Second grade was targeted for remediation of subtraction from 9 due to the presence of pervasive deficits in calculation fluency in all class sections (mean dc 2 min 17, range 0 to 54) while accuracy was high (mean accuracy 92%, range 100% to 38%). All five classrooms were selected for participation in study activities. This allowed the experimenter to work with the full population of second grade students from the rural community (to ensure a normative profile of student performance). Once the grade and skill were targeted, teacher consent was obtained before baseline data collection took place.

Baseline. During baseline all second grade students were given two minutes to complete a subtraction from 9 probe. The experimenter (primary

investigator or trained research team members) read a script to the students in each class, using the procedures outlined by standard curriculum based measurement (Shinn, 1989). An example of the script is located in the Appendix. The students were instructed not to skip problems, a deviation from Shinn's (1989) procedures, due to the group administration. If a student only attempted the easier problems on the page (for example, "1 - 1" or "1 - 0") this would not accurately reflect the student's fluency skills. The experimenter then instructed the students to begin and allowed them two minutes to complete as many problems as possible; a stopwatch or timer was used to monitor the time. At the end of the 2 minute time period the experimenter instructed the students to stop working. The experimenter then collected the probes. Within the same 24 hour period the probes were scored for digits correct per two minutes and fluency scores were entered into a database. This administration occurred once daily over the course of two days to establish stable current performance rates. Three data points for each student (screening plus two baseline points) were used to establish performance rates prior to implementation of the fluency intervention.

Accuracy versus fluency. Due to high accuracy rates among the sample used in this study (89% of the students sampled exhibited accuracy rates of 80% or higher), an acquisition intervention was not used; fluency intervention data were the primary focus of this study. Although all second grade students were included in the fluency intervention, data from students with low accuracy scores (scores averaging < 80% across baseline data points) were omitted from the

study; this resulted in the omission of data from nine students in the final database. The 80% accuracy cutoff was derived from Gickling and Thompson's (1985) research suggesting that mathematics probes should contain 70-85% known items to represent a student's instructional level for that particular task.

Intervention Phase One: Establishing Rate of Response

This phase was designed to evaluate students' response to the fluency intervention. To establish the rate or slope of response all participants were exposed to the fluency intervention once daily for twenty-four sessions. During this phase each student was administered probes consisting of "subtraction from 9" problems. These probes differed from baseline probes in that each student received an individualized probe containing his/her name and the goal for the session. The goal was derived from the previous day's performance. During the session the experimenter informed the students that if they improved upon their previous score, they would be able to pick out a reward from the reinforcement box. The experimenter then followed the same procedures used in the Baseline phase. Within the same 24 hour period the probes were scored for digits correct per two minutes and fluency scores were entered into a database. If the student exceeded the goal listed he/she would receive a sticker on his/her probe the following day (along with a new performance goal). The stickers served as a daily reinforcement, and the number of stickers was tallied each week to determine each student's access to the reinforcement box. Each Friday the students were

able to choose rewards from the reinforcement box for improving upon their scores.

Proficiency Group Identification

To ascertain the level of response to the intervention, a slope was calculated for each student. Data from the subtraction from 9 probes were scored for digits completed accurately in two minutes (dc/2min) and entered into an Excel database. These scores were graphed to display fluency growth rates for each student. A least squares regression line was then calculated for each student using fluency scores and the number of intervention sessions. The slope calculated by the regression analysis indicated each student's average increase in dc/2min across the first intervention phase.

Slope data was were used to classify students into proficiency groups. Previous research studies have used slope to differentiate response groups (Vellutino et al., 1996) and instructional levels (Burns, VanDerHeyden, & Jiban, 2006). For this study, students whose slopes were greater than one standard deviation above the mean were classified as high rate responders. This group would therefore contain the students with the strongest intervention response. Students whose slopes were within one standard deviation of the mean were classified as average rate responders. Students whose slope fell below one standard deviation from the group mean were classified as low rate responders. It was reasoned that this group would contain the students with the weakest intervention response.

Once the three proficiency groups were determined the two most proficient groups (the high and average rate responders) graduated from the study. The students in the low rate responders group were then moved to the second phase of the study.

Intervention Phase Two: Rate of Response Matching

The purpose of this phase was to increase the growth of the low rate responders to that of the average responders. To accomplish this, the same intervention outlined above was applied through all intervention phases with increasing intensity for the low rate responders group. Four students who did not respond to the initial intervention intensity continued to the response matching phase and response to an increasingly intense intervention was evaluated in a multiple baseline across subjects design.

Baseline. Because the experimental design attempted to evaluate the impact of increasing intervention intensity, response to the original intensity of the intervention was the baseline against which the increasing intensity was compared. Therefore, baseline data collection involved implementation of the intervention once per session. The intervention was implemented outside of the classroom by the experimenter and team members in the Title 1 reading room, a small classroom with a single table and chairs. The experimenter followed the exact same procedures implemented during the first intervention phase. Differences in intervention delivery involved the location of delivery (i.e. small room versus classroom setting) and the absence of classroom peers.

Intensity of two sessions daily. The same intervention used in the first intervention phase (Establishing Rate of Response Phase) was utilized through all Phase Two intervention procedures. The only difference involved intensifying intervention by increasing daily frequency to occur two times a day. The two intervention sessions occurred in one session in the title reading room and lasted approximately 8 minutes a day. A unique, randomly generated probe was utilized for each session. The performance goal for the intervention session was the highest score from the previous day's performance. If the student exceeded the goal listed he/she would receive a sticker (or two, if he/she met the goal twice) on his/her probe the following day. Stickers were tallied and reinforcements were delivered each Friday. The highest score for each session was also recorded as the dependent variable for this phase.

Students' slopes were calculated daily. The student with the lowest response rate (i.e. slope) was then moved the next intensity phase. Each student was then systematically moved to the next intervention intensity phase, even if their slopes met or exceeded the performance criterion. This was used to ensure completion of the multiple baseline data collection and to ascertain response under different intensity conditions.

Intensity of four sessions daily. Once a student finished the first phase (intervention two sessions daily), intensity was doubled to four sessions a day. The fluency intervention utilized during this condition was identical to the previous condition and implemented in the same setting. Intervention at this intensity was identical to the previous condition. The time to implement this

intervention intensity was approximately 15 minutes. A unique, randomly generated probe was utilized for each session. The performance goal (and dependent variable) for each session was the highest dcp2m of the four trials. Individual student slopes were calculated daily. Once the performance criterion was met participation was systematically discontinued for each student.

Post-Intervention Phase: Maintenance of Performance Levels

Maintenance data were collected from all second grade students ($n = 71$) to evaluate the degree to which the students maintained fluency rates post intervention. Maintenance data were collected on two sessions, the first session was approximately one month after the completion of phase two intervention and the second session was approximately two months from completion. Data collection procedures were identical to those used in the phase one intervention phase, resulting in a single probe collected during each session from each student. Performance during both maintenance points were compared for each student.

Reliability Data

Reliability of probe scores. Agreement data were collected for the fluency scores on the mathematics probes. A second experimenter rescored 25% of the math probes collected. Overall reliability of scoring was 97% (range 96 to 100) agreement on an item by item analysis.

Interobserver agreement. Interobserver agreement data were collected for both intervention phases. This involved a second experimenter holding a copy

of the intervention script (which specified six steps of the intervention) and placing a check mark next to each step that was correctly implemented; 25% of the treatment sessions were observed. The overall interobserver agreement was 94% (range 83 to 100).

Planned Statistical Analyses

Data were analyzed using a one-way analysis of variance (ANOVA). The analyses were computed using the general linear model (GLM) through SPSS software. Appropriate tests were used to ensure that all statistical assumptions were met. These tests included analysis of skewness and kurtosis and homogeneity of variance (using Levene's test). Post hoc analyses were also used. If response groups were significantly different, comparison tests were planned to examine differences between groups. A multiple baseline across subjects design was also used to examine differences in response between subjects across varying intervention intensities.

CHAPTER IV

RESULTS

Pre-Intervention screening indicated that nine participants were performing with accuracy scores of less than 80%. Although all students ($n = 82$) were included in the fluency intervention, data from students with low accuracy scores were omitted from the study. Moreover, two students were missing data (due to excessive absences) for more than 7 sessions of the study and were excluded from the analyses. Thus, there were 71 students in the total sample.

Intervention Phase One: Establishing Rate of Response

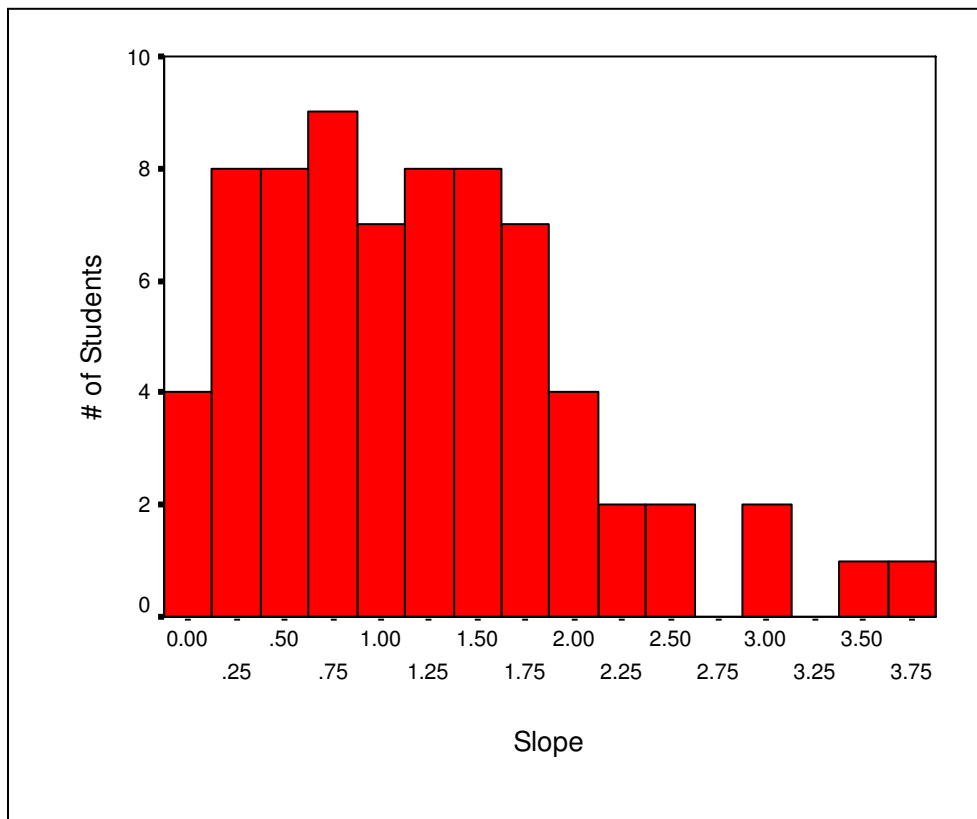
The first intervention phase was conducted for 24 sessions. When this responsiveness evaluation was completed, average performance had increased from 18 digits correct per two minutes (range 4 to 54) to a mean performance of 56 digits correct per two minutes (range 13 to 128) for all participants. Overall, the group increased their fluency scores by 311%.

Data from the probes were graphed for each individual student. Using this data, a regression line was fitted to each student based upon ordinary least squares regression using fluency scores and the number of intervention sessions. The slope calculated by the regression analysis indicated each student's average increase in $dc/2min$ across the first intervention phase. Standard error of slopes was then computed using the formula outlined by Christ

(2006). The resulting reliability was .99 for the sample, which suggested sufficiently reliable slopes for research.

To evaluate the estimated slopes' adherence to a normal distribution, the data were first analyzed by computing estimates of skewness and kurtosis. The skewness estimate was .85. The standard error of skewness was .29. Thus, a positive skew was found for the fluency slopes, meaning that the data were slightly skewed to the right of the distribution curve. The kurtosis estimate was .70, with a standard error of kurtosis of .56. This indicated that the slope data had a leptokurtic distribution (see Figure 1).

Figure 1
Distribution of Fluency Slopes for the Total Sample (N = 71)



Given the skewness and kurtosis scores of the distribution, a univariate analysis of outliers was conducted by converting the students' slope scores to z scores and then omitting those scores which had an absolute value greater than ± 3.00 . This resulted in the omission of data from two students whose slope values were greater than 3.00. The revised estimates of the new sample ($n = 69$) resulted in a skewness of 0.53 with a standard error of 0.29. The resulting kurtosis estimate was -0.11 with a standard error of 0.57 (See Figure 2). The removal of outlying data resulted in a more normal distribution. The mean slope for the final group ($n = 69$) over the 24 intervention sessions was 1.12 ($SD = 0.72$). See Table 1 for descriptive statistics of the final sample.

Figure 2
Distribution of Fluency Slopes for the Final Sample (N = 69)

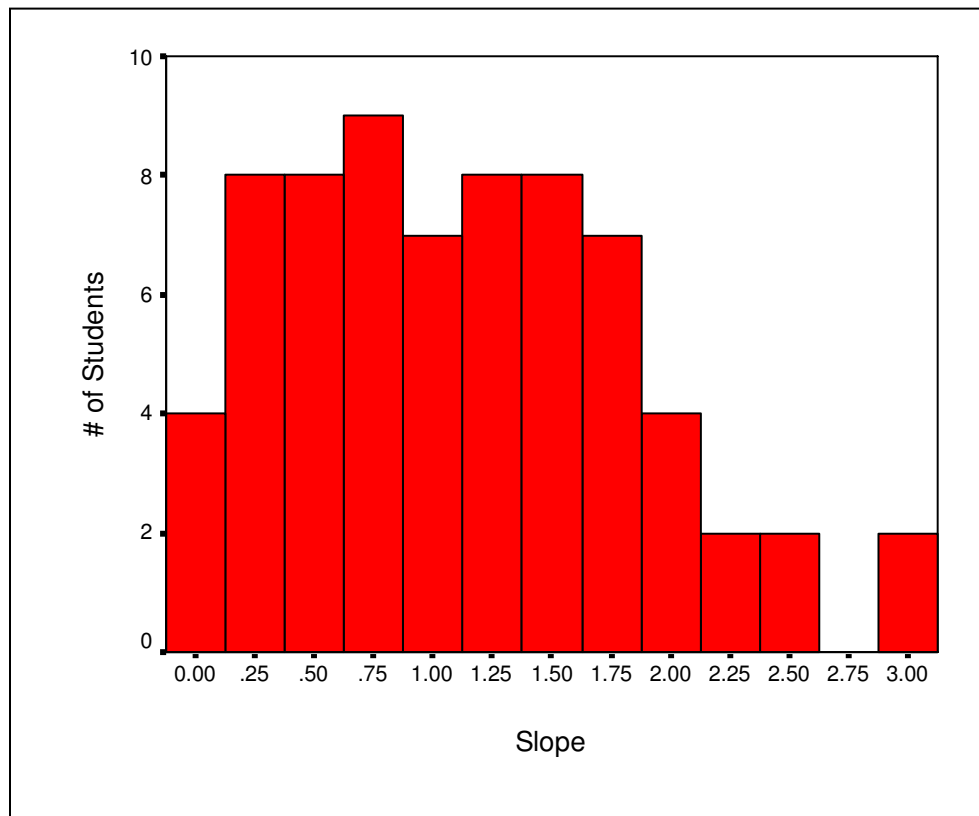


Table 1
Descriptive Statistics of the Final Sample

| | <u>N</u> | <u>Range</u> | <u>Min</u> | <u>Max</u> | <u>Mean</u> | <u>SD</u> |
|-------|----------|--------------|------------|------------|-------------|-----------|
| Slope | 69 | 3.12 | -.07 | 3.05 | 1.12 | .72 |

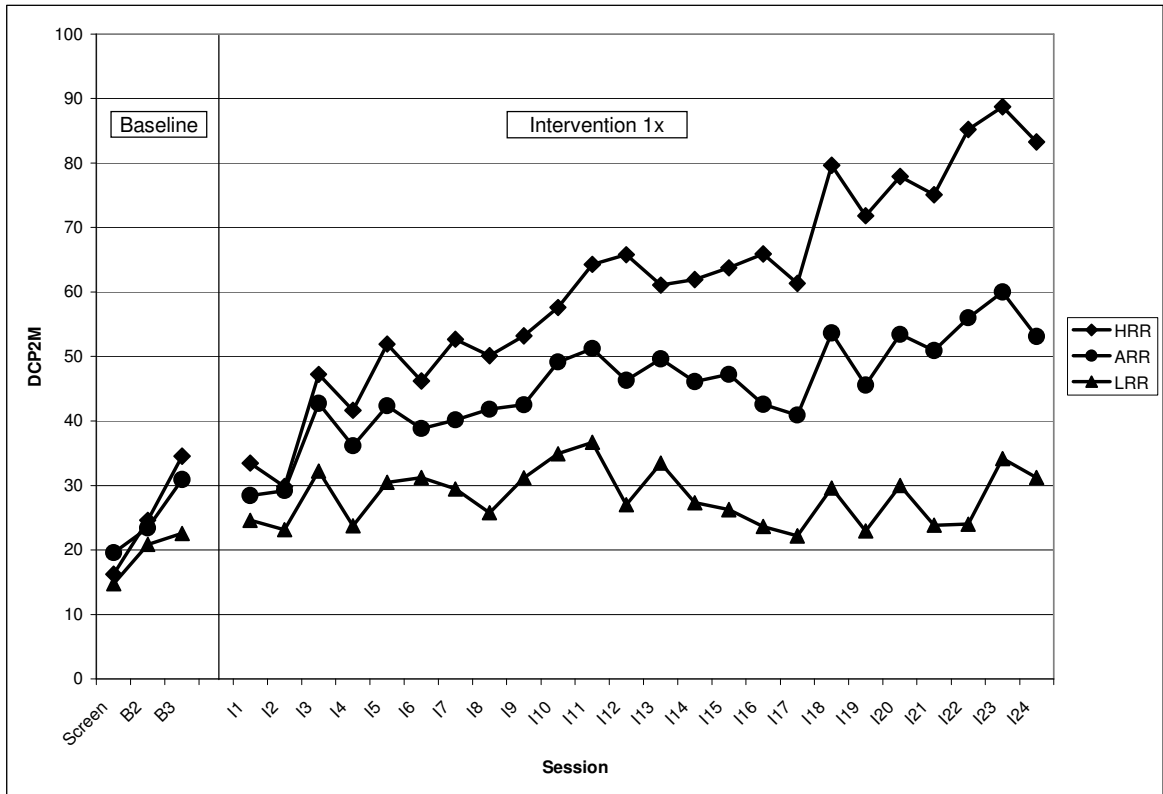
Proficiency Group Identification

Slope data were used to classify students into proficiency groups. For this study, students whose slopes were greater than one standard deviation above the mean (slope ≥ 1.84) were classified as high rate responders. The mean slope for the high rate responders group (n = 11) was 2.29 (SD = 0.41). Students whose slopes were within one standard deviation of the mean were classified as average rate responders. The mean slope for the average rate responders (n = 45) was 1.10 (SD = 0.41). Students whose slope fell below one standard deviation from the group mean (slope $\leq .40$) were classified as low rate responders. The mean slope for the low rate responders group (n = 13) was 0.20 (SD = 0.16). See Table 2 for descriptive statistics of each response group. Figure 3 displays the mean slopes for each response group.

Table 2: Descriptive Statistics for Response Groups

| | <u>N</u> | <u>M</u> | <u>SD</u> | <u>SE</u> | 95% Conf. Int. | | <u>Min</u> | <u>Max</u> |
|-------|----------|----------|-----------|-----------|----------------|--------------|------------|------------|
| | | | | | <u>Lower</u> | <u>Upper</u> | | |
| LRR | 13 | .20 | .16 | .04 | .11 | .30 | -.07 | .39 |
| ARR | 45 | 1.10 | .41 | .06 | .97 | 1.22 | .44 | 1.78 |
| HRR | 11 | 2.29 | .41 | .12 | 2.02 | 2.57 | 1.87 | 3.05 |
| Total | 69 | 1.12 | .72 | .09 | .95 | 1.29 | -.07 | 3.05 |

Figure 3
Mean Slopes of Response Groups



To determine whether the differences between the response groups were statistically significant, a one-way analysis of variance (ANOVA) was used. The groups were found to be statistically different based on slope, $F(2, 66) = 90.17, p < .001$ (see Table 3). The null hypothesis assuming homogeneity of within group variances (using Levene’s test), however, was rejected due to disparate sample sizes. To provide further support despite this failure to meet the assumption of homogeneity of variance, a nonparametric analysis, the Kruskal-Wallis test, was used to test statistical differences between the response groups. The chi-square statistic used in the Kruskal-Wallis test was significant, $\text{Chi-Square}(2) = 48.42, p < .001$, indicating that it was unlikely to obtain samples with average ranks so far apart if the null hypothesis were true. Therefore, it appears that the data are able

to confirm the original analysis because the Kruskal-Wallis test provided further evidence that population differences exist.

Table 3
Analysis of Variance of Estimated Fluency Slopes as a Function of Response Group

| Source | df | SS | MS | F |
|----------------|----|-------|-------|--------|
| Between Groups | 2 | 26.01 | 13.01 | 90.17* |
| Within Groups | 66 | 9.52 | .144 | |
| Total | 68 | 35.53 | | |

* $p < .001$

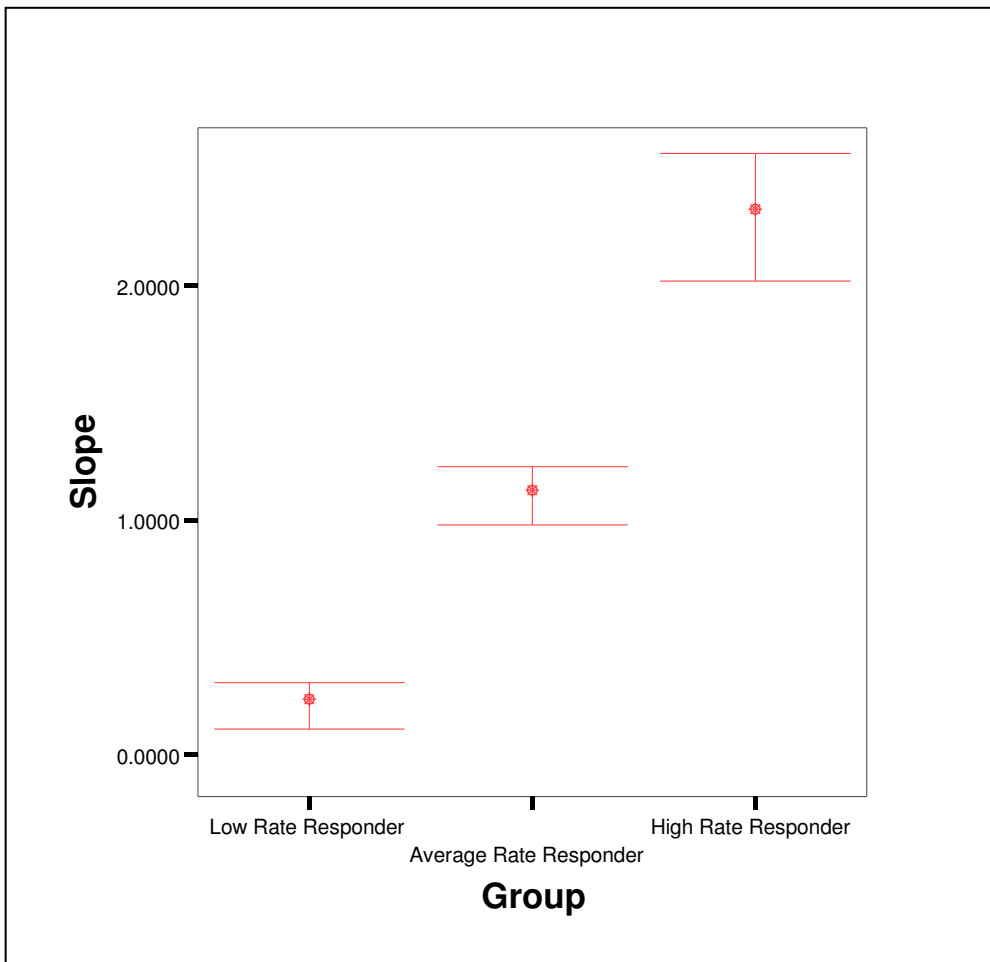
The Games-Howell multiple comparison test was used to discover which group means differed for each of the response groups. This post hoc analysis was chosen because it is robust against the effect of unequal sample sizes. The Games-Howell post hoc analysis found that all groups differed significantly from each other (see Table 4). Mean slopes from Low Rate Responders differed significantly from that of both Average Rate Responders (mean difference = $-.89$, $p < .001$) and High Rate Responders (mean difference = $-.209$, $p < .001$). Average Rate Responders were also found to differ significantly from High Rate Responders (mean difference = -1.19 , $p < .001$). Figure 4 provides a visual sense of how far the groups are separated. The confidence bands are determined for each group separately. The graph clearly presents the relation between response group and slope.

Table 4
Games-Howell Post Hoc Results

| (I) Group | (J) Group | Mean Diff | | 95% Confidence Int. | |
|-----------|-----------|-----------|-----|---------------------|-------|
| | | (I-J) | SE | Lower | Upper |
| LRR | ARR | -.89* | .08 | -1.08 | -.71 |
| | HRR | -2.09* | .13 | -2.43 | -1.74 |
| ARR | LRR | .89* | .08 | .71 | 1.08 |
| | HRR | -1.19* | .14 | -1.55 | -.84 |
| HRR | LRR | 2.09* | .13 | 1.74 | 2.44 |
| | ARR | 1.19* | .14 | .84 | 1.55 |

*p<.001

Figure 4
Error Bar Chart of Slope by Response Group



Intervention Phase Two: Rate of Response Matching

The second intervention phase (Rate of Response Matching) included 4 students: Colton, Mallory, Kenneth, and Zeke who exhibited low response rates during the first intervention phase. Each of these students was a member of the Low Rate Responders group. The slopes during the first intervention phase for Colton, Mallory, Kenneth and Zeke were 0.39, 0.35, -0.03, and -0.07, respectively.

Baseline

Performance of the four Low Rate Responders during the individual baseline (intervention once per day outside the general education classroom) was very similar to that of their slopes during the first intervention phase. Colton, Mallory, Kenneth, and Zeke continued to perform with slopes significantly lower than that of the average student in the classroom with slopes of 0.50, -0.07, 0.42, and -1.50, respectively (See Figure 4).

Intensity of Two Sessions Daily

During this phase the four participants were exposed to the intervention increased to two occurrences per day. Two of the participants' slopes exceeded the performance criterion (slope of 1.12); during this intensity phase Colton's slope increased to 3.19 and Mallory's slope increased to 1.99. Although Zeke's slope increased from -1.50 to 0.00, it did not meet the performance criterion. Kenneth's performance did not show improvement during this phase, with his slope dropping to 0.30.

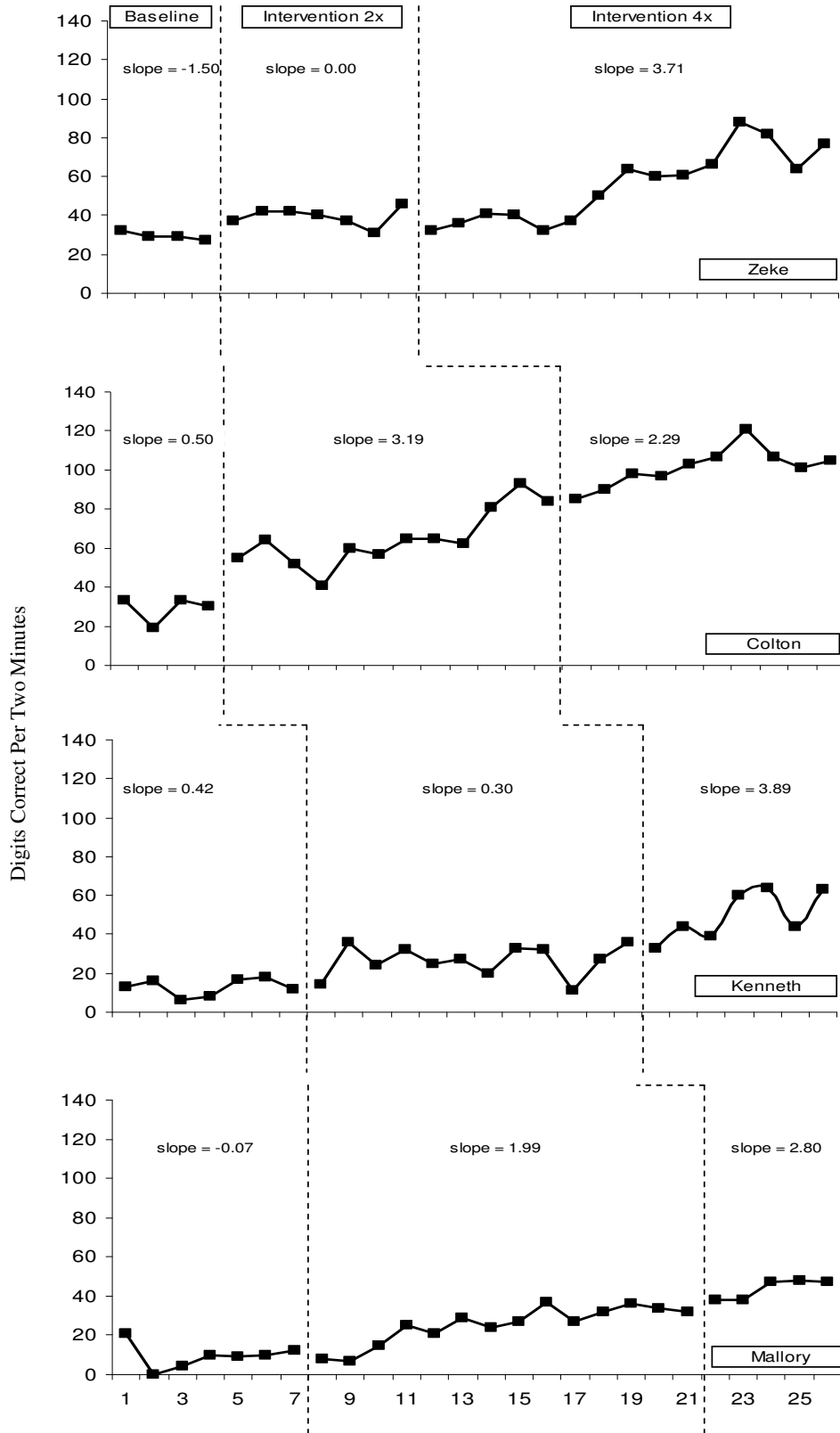
Intensity of Four Sessions Daily

During this phase the students received the intervention four times per day. During this phase, all students' slopes exceeded the performance criterion. Although they did not respond significantly to the intervention in the previous phase, both Kenneth's and Zeke's slopes exceeded that of the Average Rate Responders. Kenneth's slope increased to 3.89, while Zeke's increased to 3.71. Mallory's slope also increased during this intervention phase, climbing to 2.80. Colton's slope, however, dropped slightly, to that of 2.29. Throughout all intervention phases, all students' accuracy scores remained over 80%. Figure 5 offers a graphed interpretation of these data.

Post-Intervention Phase: Maintenance of Performance Levels

Maintenance data were collected from all second grade students to evaluate the degree to which the students maintained fluency rates post intervention. A least squares regression line was calculated for each student using his or her last two intervention (from Phase One Intervention) scores with the two additional maintenance scores. The maintenance slope for the group over the 4 sessions was 0.11 (SD = 5.49).

Figure 5
Multiple Baseline Design Across Subjects



CHAPTER V

DISCUSSION

The primary purpose of this investigation was to examine student response to a general education invention used to increase mathematics fluency. Response was examined by analyzing learning rates (i.e. slopes) of students in a second grade cohort. Using slopes allowed the researcher to quantify intervention response and gain a clearer understanding of how this response changed when intervention was increased by intensity levels (i.e. frequency).

Initially, the intervention was applied once daily to all students in the second grade cohort. Across the group, student performance increased by 311%. Slopes were analyzed and found not only to be highly reliable, but also fairly normally distributed. Analyzing slopes across all students allowed the experimenter to understand how most students would respond to a classwide intervention and assisted in establishing a baseline of “normative” response. In RTI terms, the experimenter examined student response rates which would reflect a Tier 1 intervention in the RTI literature. Once the student slopes were examined, they were then averaged to establish a metric by which to compare each student’s response.

This response metric (average slope of 1.12) was used to create proficiency groups within the distribution based on rate of response. Three

groups with significantly different average slopes were established. On average, High Rate Responders advanced two digits each intervention session. This was significantly different from the rate of performance of Average Rate Responders, who advanced one digit each session. Low Rate Responders also differed significantly from the two aforementioned groups. On average, it took approximately five sessions for the Low Rate Responders to advance by one digit. Each response group had remarkably different learning trajectories.

Results of this study indicate that the application of an intervention hierarchy design involving increased intervention frequency reduced the difference between the slope of average rate learners and low rate learners. Increased frequency intervals resulted in quantitatively more intense treatment for individual students. The use of a response metric (slope of 1.12) provided an anchor to which to compare the student's response rates. Slope rates of two of the participants (Colton and Mallory) exceeded that of the "average" student when the intervention was increased to two times its original intensity. The other two students (Zeke and Kenneth) showed insufficient response (slope < 1.12) until they were exposed to intervention intensity four times the original intervention. During this phase all students slopes increased, except for Colton's which dropped slightly. Colton's slope "flattened out" due to a likely ceiling effect, meaning he was at his maximum fluency rate for this task. On average, Colton's fluency scores were higher than the other students and it appeared that he was functioning at his highest potential. At the conclusion of the intervention phases, all students had met criterion and their response rates exceeded that of the

average response. In regards to the RTI literature, the resulting intervention would be described at a Tier 2 intervention.

Although establishing growth through intervention is necessary, maintenance of the skill after the intervention has been removed is a crucial component of intervention response. The maintenance slope was calculated by using the last two intervention scores with the two additional maintenance scores. When the maintenance screening was conducted, the resulting slope for the entire group was 0.11. This indicated that the students had maintained their previous gains from intervention. The students involved in the second part of the study also maintained their previous performance although their post-intervention gap was much shorter than those included in only the first part of the study. Knowledge of maintenance of skills obtained through intervention is important when making decisions about service delivery. Practitioners need to examine not only the amount of intervention needed to produce performance, but the level required to maintain that performance as well. This knowledge can help address questions concerning the type of programming necessary to help remediate academic deficits.

Implications for Practice

A purpose of this study was to understand what “normal” intervention response looks like amongst a cohort of second grade students. Understanding how students perform, on average, regarding a particular academic skill allows a basis for comparison. Making decisions regarding general versus special

education placement is more objective when there is a documented, analytic rate of comparison. The need for a normative response rate also dictates the need for comprehensive screening and thorough progress monitoring data; both should be collected at multiple times throughout the academic year.

Another purpose of this study was to quantify individual response rates to establish a metric of response. This metric was then used to compare response rates across different quantities of intervention. The combination allowed the experimenter to objectively assess student response rates and need for increased intervention frequency. Collecting data in this manner may help decisions regarding special education services. For example, if a student requires five times the amount of an intervention as an average student to make and sustain educational gains, then it is possible that the student may require programming in the special education environment. Researchers, along with practitioners and administrative staff need to determine what level of intervention is too much for the general education environment. Studies such as this may help to make that decision-making process more objective because it details how much intervention a student needs to make gains regardless of academic placement.

A third purpose of this study was to provide staff members (within this school district) with information regarding the intensity of intervention necessary to produce growth in the majority of their students. Knowledge regarding the amount of intervention necessary to improve student performance is crucial for effective and efficient use of resources, especially in financially strained public

schools. Data from this study indicated that implementing this particular intervention (goal setting and reinforced practice) four sessions a day improved the majority of students' performance. Teachers were provided with this information, not only to intervene when students are indicating low fluency performance regarding this skill, but also as a means of prevention for future cohorts.

Limitations and Directions for Future Research

Although the results of this study procured positive outcomes for the students involved and provided insight on how to quantify intervention response, there are several limitations that should be taken into account when analyzing these results. The first limitation involves the fact that the data were collected within one school district from a rural area. This may cause the generalizability of the data to be in question, although there were no obvious components of the intervention that would suggest differences in utility across populations. Another limitation involved the collection of only one form of data (i.e. probe fluency data). Collecting a second form of fluency data, possibly from a more standardized measure, would have helped to establish a more generalizable benefit from this intervention.

Another limitation of the study involved the focus on fluency data, as opposed to the examination of both accuracy and its relation to fluency. Data from students whose accuracy scores were less than 80% were excluded from the

study. Since adequate math fluency cannot be gained without accuracy, a more comprehensive study may rate the gains in both accuracy and fluency.

A final limitation of the study involves the extension to other academic domains. This study only examined simple subtraction fluency response and therefore its application to other areas in mathematics is questionable. Future researchers may wish to replicate the current study with other mathematic domains.

There is a continued need to examine response to intervention models in multiple academic areas such as mathematics (Vaugh & Fuchs, 2003). There is also a need to further understand the level of intervention necessary to produce adequate response levels. For practical purposes, it is necessary to understand the amount of intervention that can produce the most beneficial outcomes. For example, in the second part of the study Colton only required the intervention twice that of its original intensity to produce adequate intervention response. When provided the intervention four times a day his response did not increase. When a researcher or practitioner finds an intervention is no longer beneficial to the student, he/she must consider three things: 1. Does the intervention need to be changed? 2. Does the intervention need to be intensified in some way; or 3. Is the student performing at his/her highest capacity with this skill? The third question involves a concept called diminishing returns. There may come a point when the intensity of intervention does not justify the amount of return. More research in this area is necessary. Researchers and practitioners need to discover how to best produce meaningful results for students in a way that is

most reasonable given the amount of resources in financially strained public schools.

REFERENCES

- Aaron, P. G. (1997). The impending demise of the discrepancy formula. *Review of Educational Research, 67*, 461-502.
- Algozzine, B., & Ysseldyke, J. E. (1986). The future of the LD field: Screening and diagnosis. *Journal of Learning Disabilities, 19* (7), 394-398.
- Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, F. E. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *The Journal of Special Education, 38* (2), 66-79.
- Burns, M. K., & Senesac, B.V. (2005). Comparison of dual discrepancy criteria to assess response to intervention. *Journal of School Psychology, 43*, 393-406.
- Burns, M. S., VanDerHeyden, A. M., & Jiban, C. L. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35* (3), 401-418.
- Case, L. P., Speece, D. L., & Molloy, D. E. (2003). The validity of a response-to-instruction paradigm to identify reading disabilities: A longitudinal analysis of individual differences and contextual factors. *School Psychology Review, 32*(4), 557-582.
- Christ, T. J. (2006). Short-term estimates of growth using curriculum-based

measurement of oral reading fluency: Estimating standard error of slope to construct confidence intervals. *School Psychology Review*, 35 (1), 128-133.

Danielson, L., Doolittle, J., & Bradley, R. (2005). Past accomplishments and future challenges. *Learning Disability Quarterly*, 28, 137-139.

Dean, V. J., & Burns, M. K. (2002). Inclusion of intrinsic processing difficulties in LD diagnostic models: A critical review. *Learning Disability Quarterly*, 25, 170-176.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37 (3), 184-192.

Deno, S. L., Fuchs, L. S., Marston, D., & Shinn, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30 (4), 507-524.

Dombowski, S. C. (2003). Norm-referenced versus curriculum-based measurement: A balanced perspective. *NASP Communique*, 31 (7), 1-12.

Freeland, J. T., & Noell, G. H. (1999). Maintaining accurate math responses in elementary school students: The effect of delayed intermittent reinforcement and programming common stimuli. *Journal of Applied Behavior Analysis*, 32 (2), 211-215.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, 27, 216-227.

Fuchs, D., Roberts, P. H., Fuchs, L. S., & Bowers, J. (1996). Reintegrating

- students with learning disabilities into the mainstream: A two-year study. *Learning Disabilities Research and Practice, 11*, 214-229.
- Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research and Practice, 18* (3), 172-186.
- Fuchs, L. S., Bahr, C. M., & Rieth, H. J. (1989). Effects of goal structures and performance contingencies on math performance of adolescents with learning disabilities. *Journal of Learning Disabilities, 22*, 554-560.
- Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204-219.
- Fuchs, L.S., & Fuchs, D. (2005). Enhancing mathematical problem solving for students with disabilities. *The Journal of Special Education, 39* (1), 45-57.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. I. (1989). Effects of alternative goal structures within curriculum-based measurement. *Exceptional Children, 55*, 429-438.
- Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*, 33-45.
- Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18* (3), 157-171.

- Gickling, E., & Thompson, V. (1985). A personal view of curriculum-based assessment. *Exceptional Children, 52*, 205-218.
- Gresham, F. M. (2001). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. *Paper prepared for the OSEP Learning Disabilities Initiative, Office of Special Education Programs, U.S. Department of Education: Washington, DC.*
- Gresham, F. M., VanDerHeyden, A., & Witt, J. C. (2005). Response to intervention in the identification of learning disabilities: Empirical support and future challenges. *Paper prepared for the OSEP Learning Disabilities Initiative, Office of Special Education Programs, U.S. Department of Education: Washington, DC.*
- Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology, 38*, 25-33.
- Lentz, F. E., Jr., Allen, S. J., & Ehrhardt, K. E. (1996). The conceptual elements of strong interventions in school settings. *School Psychology Quarterly, 11*, 118-136.
- Lindamood, P., & Lindamood, P. (1998). *The Lindamood phoneme sequencing program for reading, spelling, and speech*. Austin, TX: PRO-ED
- Lyon, G. R. (1996). Learning disabilities. *The Future of Children: Special Education for Students With Disabilities, 6 (1)*, 54-76.
- Marston, D., Fuchs, L. S., & Deno, S. I. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related

- measures. *Diagnostique, 11*, 71-90.
- Mastropieri, M. A., Scruggs, T. E., & Shiah, S. (1991). Mathematics instruction for learning disabled students: A review of research. *Learning Disabilities Research and Practice, 6*, 89-98.
- McConnell, S. R., Rush, K. L., McEvoy, M. A., Carta, J., Atwater, J., & Williams, R. (2002). Descriptive and experimental analysis of child-caregiver interactions that promote development of young children exposed parentally to drugs and alcohol. *Journal of Behavioral Education, 11*, 131-161.
- Noell, G. H., Duhon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review, 31* (2), 217-234.
- Mortenson, B. P., & Witt, J. C. (1998). The use of weekly performance feedback to increase teacher implementation of a prereferral academic intervention. *School Psychologist, 50*, 741-749.
- Oehler-Stinnett, J. J. (1986). The ability of the Luria-Nebraska Neuropsychological Battery--Children's Revision to discriminate between learning disabled and slow learner children. *Unpublished Doctoral Dissertation*. University of Southern Mississippi, Hattiesburg, Mississippi.
- Proctor, B., & Prevatt, F. (2003). Agreement among four models used for diagnosing learning disabilities. *Journal of Learning Disabilities, 36*(5), 459-466.
- Rhymer, K. N., Dittmer, K. I., Skinner, C. H., & Jackson, B. (2000). Effectiveness

of a multi-component treatment for improving mathematics fluency.

School Psychology Quarterly, 15, 40-51.

Rutter, M. (1989). Isle of Wight revisited: Twenty-five years of child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 633-653.

Schunk, D. H. (1985). Participation in goal setting: Effects on self-efficacy and skills of learning disabled children. *The Journal of Special Education*, 19, 307-317.

Sheridan, S. M., Kratochwill, T. R., & Elliott, S. N. (1990). Behavioral consultation with parents and teachers: Delivering treatment for socially withdrawn children at home and at school. *School Psychology Review*: 19, 33-52.

Shinn, M R. (1989). *Curriculum-based measurement: Assessing special children*. (2nd ed.). New York: Guilford.

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.

Smith, D. D., & Lovitt, T. C. (1976). The differential effects of reinforcement contingencies on arithmetic performance. *Journal of Learning Disabilities*, 9, 32-40.

Sofie, C. A., & Riccio, C. A. (2002). A comparison of multiple methods for the identification of children with reading disabilities. *Journal of Learning Disabilities*, 35 (3), 234-244.

- Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology, 93* (4), 735-749.
- Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice, 18* (3), 147-156.
- Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review, 29* (3), 443-461.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34* (1), 33-58.
- Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice, 18*, 137-146.
- 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., Sipay, E., Small, S., Pratt, A., Chen, R., & Denckla, M. (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, 601-638.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: American Guidance Services.

Ysseldyke, J. (2005). Assessment and decision making for students with learning disabilities: What if this is as good as it gets? *Learning Disability Quarterly*, 28, 125-128.

Ysseldyke, J. (2001). Reflections on a career: 25 years of research on assessment and instructional decision-making. *Exceptional Children*, 67 (3), 295-310.

APPENDICES

Appendix A *Research Prospectus*

Research Project Synopsis

Title: Response to Intervention: Incorporation of an Increasing Intensity Design to Improve Mathematics Fluency

Investigators:

Michelle Atkins, M.A.—Doctoral Student, School Psychology Program
Gary Duhon, Ph.D.—Oklahoma State University; School Psychology Program

Purpose of Research:

With the reauthorization of the Individuals with Disabilities Education Act (IDEA), there has been much discussion as to which model should be used to identify learning disabilities. As criticism of the discrepancy-based approach intensifies, the new IDEA regulations have suggested the adoption of a response to intervention approach. Response to intervention is defined as a change in performance or behavior due to intervention (Fuchs, Mock, Morgan, & Young, 2003; Gresham, 1991; Gresham, 2001). This framework was originally envisioned by Heller, Holtzman, and Messick in 1982 (Fuchs, 2003) and was operationalized by Fuchs and Fuchs in 1995 (Vellutino, Scanlon, Sipay, Small, Pratt, Chen, et al., 1996). This approach is based on discrepancy, but the discrepancy is between pre- and post-intervention levels instead of ability and achievement scores. This model deals with a *lack* of discrepancy since the goal of intervention is to produce a difference between pre- and post-test scores.

One function of this approach is to lead to better classroom interventions. It focuses on maximizing the effectiveness of regular education for all students. Although this approach seems promising, more research is necessary in order to examine its utility and feasibility in the classroom.

The purpose of the intervention is to improve fluency in a keystone skill that is an essential component of the curriculum. This project will evaluate the effectiveness of a

particular intervention used to improve fluency in single-digit multiplication. It will also evaluate the amount of intervention necessary to help improve all students' functioning in this domain.

Specific Objectives:

This research project is designed to evaluate two areas. (1) This study seeks to evaluate the degree to which students respond to a simple intervention found to be effective in previous (Mastropieri, Scruggs, & Shiah, 1991). (2) This study also seeks to evaluate the amount of intervention that will be the most effective for all students. After the study, this information would be disseminated to both the principal and teachers regarding what was found to be the most effective and *efficient* intervention for students to make gains quickly with minimal use of resources.

Instrumentation/Materials:

Materials for this study will consist of mathematics probes comprised of single digit math problems. This skill would be identified Priority Academic Student Skill (Oklahoma PASS Standards).

Reinforcements will be used during the intervention phases of the study in an attempt to ensure high levels of effort throughout the study. A reinforcement survey will be distributed to the third grade teachers (including small school supply items such as pencils, stickers, etc. for them to choose from) in order to identify acceptable reinforcers for the classroom. A reinforcement box will be created using information from the survey.

Target Population:

The participants in this study will include students and general education teachers from XXXX Public Schools. Participants will be students from three grade classrooms. Students will be given permission forms which must be signed by their parents in order to be included in the study. After receiving parent permission, child assent will also be obtained. As stated in both parent permission and child assent forms the student can withdraw permission at any time to remove themselves from the research project.

Research Conditions:

Phase one: The first part of this project will involve a class-wide intervention conducted in the general education classroom. During this phase each student will be administered probes consisting of single digit math problems. Each student will receive an individualized probe containing their name and the goal for the session. The goal will derive from the previous day's performance and the students will be told that if they

improve upon their previous score, they will be able to pick out a reward from the reinforcement box. The students will have two minutes to work on the math probe. If the student exceeds the goal listed he/she will receive a checkmark for that day. At the end of each week, the experimenter will bring a chart listing how many times each student has improved upon his or her score. Each student will then have access to the reinforcement box and will be able to choose a number of rewards (based on the amount of times he/she improved upon his/her score). The amount of time the intervention will take each day is 5 minutes.

Phase two: The second part of the project will consist of a small-group intervention and will be run in the guidance, enrichment, or music room depending on the day and availability. The purpose of this phase is to increase the growth of students with lower fluency scores in the previous phase. To accomplish this, the same intervention outlined above will be applied through all intervention phases with increasing intensity for the low rate responders group. Initially, the intervention intensity will double, meaning that the intervention will be administered on two occasions during the daily session. The performance of the students will be evaluated in order to ensure that they are making adequate progress. If little to no growth is made, then the intervention will double again, meaning that the intervention will be administered on four occasions during the daily session. Intervention intensity will continue to increase (double) until each student is making adequate growth. The amount of time this intervention will take each day is 10 to 15 minutes.

Confidentiality Procedures:

A database will be set up using data gathered from this study which will contain teacher and student names and scores obtained from the math probes. However, this database is contained within a password-protected program with access only available to the researchers working on this project. Any data reported to the general public would be group data. Individual scores would not be disseminated. No identifying information (student, teacher, school, district) will be made public.

At the end of the study the teachers will be given information concerning their students' performance. Parents who request information regarding their child's progress will also receive information concerning their multiplication performance. The school principal will also receive overall data concerning students' scores.

Utilization of Results

The data collected from this study will be used for the purposes of completing a dissertation and publication of said dissertation. The results of this study will also be used for program planning for children in need of intervention services.

Appendix B
Single Skill Probe

Name: Jane Doe
Date: May 1, 2008

Class: Smith

Last time your score was 25. Good Job! Try to beat your score!
Subtraction from 9

$$\begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 6 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array}$$

$$\begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array}$$

$$\begin{array}{r} 5 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 6 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array}$$

$$\begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 8 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 7 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 4 \\ \hline \end{array}$$

$$\begin{array}{r} 9 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 9 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 6 \\ \hline \end{array}$$

$$\begin{array}{r} 7 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 7 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array}$$

$$\begin{array}{r} 6 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 6 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 9 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 3 \\ \hline \end{array}$$

$$\begin{array}{r} 8 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 6 \\ \hline \end{array} \quad \begin{array}{r} 4 \\ - 4 \\ \hline \end{array} \quad \begin{array}{r} 3 \\ - 3 \\ \hline \end{array} \quad \begin{array}{r} 5 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 6 \\ - 5 \\ \hline \end{array} \quad \begin{array}{r} 8 \\ - 7 \\ \hline \end{array}$$

Appendix C
IRB Approval Letter

Oklahoma State University Institutional Review Board

Date: Thursday, September 28, 2006
IRB Application No ED06173
Proposal Title: Response to Intervention: Incorporation of an Increasing Intensity Design to Improve Mathematics Fluency

Reviewed and Processed as: Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 9/27/2007

Principal Investigator(s)

Michelle Atkins
434 Willard
Stillwater, OK 74078

Gary J Duhon
423 Willard
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Beth McTernan in 219 Cordell North (phone: 405-744-5700, beth.mcternan@okstate.edu).

Sincerely,



Sue C. Jacobs, Chair
Institutional Review Board

VITA

Michelle Elizabeth Atkins

Candidate for the Degree of

Doctor of Philosophy

Report: RESPONSE TO INTERVENTION: INCORPORATION OF AN INCREASING INTENSITY DESIGN TO IMPROVE MATHEMATICS FLUENCY

Major Field: School Psychology

Biographical:

Education: Graduated from Oak Ridge High School, Conroe, Texas in May 1996. Received Bachelor of Arts degree in Psychology from University of St. Thomas, Houston, Texas in May 2000. Received Masters of Arts degree in Developmental Psychology from University of Houston, Houston, Texas in August 2003. Completed the Requirements for the Doctor of Philosophy in Educational Psychology (Option: School Psychology) degree at Oklahoma State University, Stillwater, Oklahoma in July 2008.

Experience: University of St. Thomas Undergraduate Research Assistant, December 1998 – May 1999; McBride Institute, College Bound Volunteer, January 1999 – May 1999; Kidventure, Camp Counselor, May 1999 – August 1999; University of Houston, Health Psychology Research Group Member, August 2000 – August 2003; University of Houston, Teaching Fellow, August 2002 – July 2003; Oklahoma State University, Course Instructor, August 2003 – May 2005; Oklahoma State University, Research Team Member; September 2003 – May 2007; Cypress Fairbanks ISD APPIC Pre-Doctoral Internship, Intern, August 2007 – June 2008.

Professional Memberships: National Association of School Psychologists; Texas Association of School Psychologists; Oklahoma School Psychological Association.

Name: Michelle Atkins

Date of Degree: July, 2008

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: RESPONSE TO INTERVENTION: INCORPORATION OF AN
 INCREASING INTENSITY DESIGN TO IMPROVE
 MATHEMATICS FLUENCY

Pages in Study: 89

Candidate for the Degree of Doctor of Philosophy

Major Field: School Psychology

Although there is ample evidence regarding the effectiveness of the response to intervention (RTI) model for assessing reading disabilities, the utility of the RTI model in assessing other academic domains has remained relatively unexplored. The purpose of this study was to investigate an RTI model in mathematics wherein an effective general education intervention was implemented with students in the second grade. Participants in this study were 5 teachers and 71 general education students from an elementary school in a Southwest rural community. The first phase of this study involved a class-wide intervention utilizing goal-setting and reinforced practice and was implemented over 24 sessions. Slopes were calculated for each student using least squares regression to determine learning trajectories. These slopes were then used to establish proficiency groups (High Rate Responders, Average Rate Responders, and Low Rate Responders) based on the students' response to intervention. Each response group had remarkably different learning trajectories. Low Rate Responders ($n = 4$) were included in the second phase of the study. Phase two entailed using the same intervention at increasing intensity levels to improve the response (i.e. slope) of students who responded poorly to the initial intervention phase. The students' response to increasing intervention intensity was evaluated within a multiple baseline design across subjects. Results of this study indicate that the application of an intervention hierarchy design involving increased intervention frequency reduced the difference between the slope of average rate learners and low rate learners.

ADVISER'S APPROVAL: Dr. Gary Duhon
