A CASE STUDY OF RESPONSIVENESS-TO-INTERVENTION TIER II TUTORING INTERVENTION: WHAT MAKES RESPONDERS AND NON-RESPONDERS DIFFERENT?

By

Wen Juan Zhang

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements for

the degree of

MASTER OF SCIENCE

in

Special Education

December, 2011

Nashville, Tennessee

Approved:

Professor Douglas H. Fuchs

Professor Lynn S. Fuchs

TABLE OF CONTENTS

LIST OF TALBES

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUTION

Responsiveness-to-Intervention (RTI), favored in the most recent reauthorization of Individuals with Disabilities Education Improvement Act in 2004 (P.L. 108-446), has been replacing the "discrepancy" between IQ and achievement approach in identification of learning disabilities (LD) (Fuchs, Compton, Fuchs, Bryant, & Davis, 2008).With the discrepancy approach, students who are struggling in regular classrooms have to wait until their academic performance fall behind significantly before they are provided with more appropriate placement and instruction. Another limitation of the discrepancy approach is that it serves solely as an index about how far the student has fallen behind, it provides little or no information regarding what kind of intervention or instruction is appropriate for a particular child. RTI, on the contrary, is an on-going process of early identification, assessment, placement, intervention, and evaluation in helping children with LD. Scientifically-validated and evidence-based intervention is the focus of RTI. The purpose of RTI is not solely "detecting" students with LD, but also preventing them from developing LD by early identification and intervening with effective instruction.

A multi-tiered intervention system is employed in all RTI models (Fuchs & Fuchs, 2006). The number of tiers may vary in different models, but the intensity of instruction progressively increases in each higher tier in all RTI models. General classroom education represents first tier or primary intervention. The majority of students responds to regular classroom instruction and makes adequate progress. Those who do not make satisfactory progress with primary

intervention move to the second tier. These students are pulled out of their regular classrooms to receive evidence-based effective instruction in small groups or in one-on-one settings. The nature of the secondary intervention is more intensive, more explicit and more teacher-directed than that of the primary intervention. Because the instruction used in the second tier is a standard scientifically validated protocol, the students who do not respond are believed to have the greatest difficulties in academic performance even with the provision of evidence-based effective instruction. Thus, they become candidates for the most intensive intervention, a higher or the highest tier - special education. RTI models are dynamic systems. Students move from one tier to another according to their responsiveness to the instruction in each tier.

The secondary intervention is the most important element in RTI because non-responders in this tier are likely to be identified with LD and responders are given early intervention to prevent from LD. Students who repeatedly fail to respond in the second tier should receive individualized special education instruction. Research evidence shows that even with the most effective instruction, there is always a small group of students who are non-responsive (e. g., Al Otaiba & Fuchs, 2006). What makes the responders and the non-responders in secondary intervention different? What characteristics help the responders to make progress? Do the non-responders share the same traits as well? These questions still need to be answered by researchers.

The purpose of this study is to evaluate the effects of a RTI Tier II decoding fluency tutoring program on first grade students who are at risk for reading disabilities. Another goal is to explore the differences between responders and non-responders and the cause of these differences. I am a research assistant who participated in a secondary intervention study in middle Tennessee for two years, tutoring 10 students (five each year), three times a week with

each student. The study is a small portion of a large-scale, longitudinal experimental RTI research project funded by the National Institutes of Health. This article discusses the students I have worked with in depth.

CHAPTER II

RESEARCH METHODS

Participants

The participants in the current study are part of a large randomized control experimental research project (PI: Doug Fuchs). The participants in Dr. Fuchs' study are selected through the following steps:

Step 1. The project coordinator recruited teachers in the Metropolitan Nashville Public Schools to participate in the study after receiving permission from the principals.

Step 2. The teachers who agreed to participate in the study selected the lowest performing students in their class. Parent Consent forms were sent to these students' caregivers.

Step 3. The students whose parents consented to participate in the study were screened. In Year 1 (school year 2009-2010), 279 students were screened. In Year 2 (school year 2010-2011), 613 students were screened. The following testing batteries were used for screening: a) Fuchs Rapid Letter Naming Test; b) Fuchs Rapid Sound Naming Test; c) Test of Word Reading Efficiency; d) Sight Word Efficiency subtest; e) Test of Word Reading Efficiency; f) Phonemic Decoding Efficiency subtest; g) Fuchs Word Identification Fluency, Form A, Passages 1 and 2; h) Woodcock Reading Mastery Test, Revised, Word Identification subtest; and i) Woodcock Reading Mastery Test, Revised, Word Attack subtest. These screening measures were combined

into a single factor score. The lowest performing students (243 in Year 1 and 349 in Year 2) among all screened students were included in testing battery 1. In year 1, all 243 students selected after screening measures participated in the study. In year 2, because the participant pool was larger than Year 1, and the research staff could only serve 260 students, additional criteria were used to eliminate some participants. Out of the 349 students who were selected after screening batteries in Year 2, 343 were tested for Battery 1. Battery 1 includes a) Wechsler Abbreviated Scale of Intelligence, Vocabulary subtest; b) Wechsler Abbreviated Scale of Intelligence, Matrix Reasoning subtest; and c) Woodcock-Johnson Tests of Achievement, Passage Comprehension subtest. WASI Vocabulary and Matrix Reasoning *T*-scores were calculated from the Battery 1 WASI scores. To be included in the study, students had to have *T*-scores greater than or equal to 37 on either of the two tests. Students' English proficiency information was collected to assure that students could understand tutoring. To be included, ELL (English Language Learner) students had to have Early Advanced or greater proficiency at the end of Kindergarten. Students' schedule (availability for tutoring) was also considered for participant selection. 260 students met the criteria and were included in Year 2. The participants in Year 2 were lower performers in reading than those in Year 1 because of the participant selection procedures.

Step 4. The target students were then randomly assigned to control (business-as-usual) group or treatment (tutoring) group.

I am one of the 40 research staff members in Dr. Fuchs' study. I was assigned to tutor five students in Year 1 and another five in Year 2.

All students in Year 1 were boys. Among them three were African American, one Hispanic and one Caucasian. Four students had free or reduced lunch. One had an Individualized

Educational Plan (IEP) with a diagnosis of LD. And two had been retained before. Three students were six years old at the beginning of the study, and two were seven years old.

In Year 2, three were three boys and two girls. Among them two were African American, and three were Hispanic. All students had free or reduced lunch. One had an Individualized Educational Plan (IEP) with a diagnosis of LD. No one had been retained before. All students were six years old at the beginning of the study.

Settings

The intervention was conducted in unused classrooms/offices/libraries at Metropolitan Nashville Public schools. These rooms can be any quiet rooms at a school with a table/desk, and at least two chairs. During the intervention, one student and I were alone in the room and we were seated next to each other at a table/desk. If these places were not available, the intervention was conducted in school hallways where possible distractions may exist. For students with attention problems, I always used a quiet place to tutor instead of a hallway.

Materials

The materials include a "Taylor" timer, a digital recorder, a student binder with lesson pages (See Appendix A for a lesson page example), pencils, point sheets, story books, incentives, and a tutoring script developed by Doug Fuchs and his research team.

Treatment

In Year 1, tutoring was conducted 3 sessions a week for 18 weeks (total number of sessions = 54). In Year 2, tutoring was conducted 3 sessions a week for 20 weeks (total number of sessions = 60). Each session lasted approximately 30 minutes.

The tutoring script was based on the secondary prevention PALS program, which has been reported to be effective with a large population of young children (Fuchs & Fuchs, 2005). The PALS program has been revised to better serve first-grade students who are at risk for RD. The script consists 25 minutes of decoding and 5 minutes of fluency. There are 5 activities on decoding: (a) sight words; (b) letter-sound correspondence; (c) decodable words; (d) story words; and (e) story reading (short stories/passages with sight words and decodable words in the lesson). Fluency follows right after the last activity (i.e., the tutor reads the story and the student repeats) in decoding. The students are asked to read the story/passage as fast as they can for 30 seconds. The tutor marks the last word the students read. The students have two chances to read more words in additional 30 seconds. If the students finish reading the story/passage in less than 30 seconds, the goal is to spend even less time on the next try. If the students read further or faster on the second or third try, a sticker is awarded.

Treatment Fidelity

All tutoring sessions are scripted and easy to follow, which ensures consistency among all students. All sessions are also interactive, which permits the delivery of the tutoring in a

natural learning style, instead of simply reading the scripts to the students. In both Year 1 and Year 2, I received extensive tutoring training, which included (a) lecturing and modeling by the project coordinators, (b) role playing with other research assistants for approximately 10 hours, and (c) independent practice for approximately 10 hours. The author also attended a staff meeting one hour each week. The majority of the meeting time was spent on addressing tutoring questions. Before working with the students, I conducted a tutoring session with a project coordinator. Corrective feedback was provided.

A checklist was used to measure fidelity. The checklist includes key information in the script. Every session was also audiotaped. In both Year 1 and Year 2, three project coordinators observed 4 live tutoring sessions the author conducted. Two research assistants independently coded 20% of the author's digital files against the checklist. Treatment Fidelity and intercoder agreement were both ≥90% on average.

Measures

In both Year 1 and Year 2, the author received extensive testing training, which included (a) lecturing and modeling by the project coordinator(s) for approximately 10 hours, (b) role playing with other research assistants for approximately 120 hours, and (c) independent practice for approximately 20 hours. After each test, the author's scoring was double checked by a project coordinator or a research assistant. Audiotapes were checked if the second scorer had any questions. After all of the data had been re-scored and checked for errors, two research assistants entered the data independently to ensure accuracy. If there was a discrepancy, a project

coordinator rechecked the original test protocols and corrected the database until no discrepancies were found.

Total agreement was used to calculate inter-rater reliability for 20% of the testing sessions. Total agreement is the most appropriate statistic to calculate inter-rater reliability for event product data (Kennedy, 2005). The formula for this calculation is (total number of agreement) ÷ (total number of agreement+ total number of disagreement) × 100%. For example, if two research assistants score the same on 28 items and differently on 2 items, the inter-rater reliability will be 28 ÷ (28+2) × 100% = 93%. The inter-rater reliability on all measures were ≥ 90%.

The following measures were conducted in this study:

Demographic information

The teachers completed two student demographic forms. The first demographic form was collected at the beginning of the school year. It included information about students' date of birth, sex, ELL, free and reduced lunch, race, and whether they were retained previously. The second form was collected near the end of the school year. It asked about students' number of absences, and whether the teachers planned or recommended to retain the students for the next school year. Teachers were also asked to report the students' IEPs. This form asked whether the student had a reading disability, math disability, orthopedic impairment, deafness/hearing impairment, blindness/visual impairment, traumatic brain injury, mental retardation, autism, speed or language impairment, emotional disturbance, ADD/ADHD, Other Health Impairment, or Other.

Student attention information

The SWAN is a rating scale developed by James M. Swanson, Ph.D. at University of California, Irvine to measure students' abilities to focus attention, control activity and inhibit impulses (Swanson et al., 2001). Both the teachers and I completed the survey at the beginning and the end of the tutoring treatment. The target students were compared to average children of the same age on 18 items using a scale of 1-7 (i.e., 1= far below, 2 =below, 3 = slightly below, 4 = average, 5 =slightly above, 6 =above, and 7= far above).

Reading achievement

The Word Identification (Word ID) subtest and the Word Attack subtest of the Woodcock Reading Mastery Test—Revised (Woodcock, 1998) were administered both before and after treatment implementation to measure the students' change in real-word and non-word reading ability. Word ID measures phonetic reading ability with sight words. There are a total of 106 sight words in this test. The difficulty level increases as the test goes on. The students were asked to read the words out of context. One point is recorded for correct answers and 0 for incorrect answers or no responses. The test begins with item 1 and stops after the student answers six consecutive test items incorrectly. The student's score is the total number of correct items.

The Word Attack test measures phonetic reading ability with made-up words. There are a total of 45 made-up words in this test. The words are also arranged in order of difficulty level, starting with the easiest words and ending with the most difficult ones. Unlike Word ID, there were two practice items that were used to train students in Word Attack. The test begins with

item 1 and stops after the student answers six consecutive test items incorrectly. One point is recorded for correct answers and 0 for incorrect answers or no responses. The student's score is the total number of correct items.

Progress Monitoring

Curriculum-based measurement word identification fluency (WIF) was used to assess on-going short-term reading progress and reading fluency. WIF was reported to be moderately or highly correlated to reading fluency and reading comprehension (Fuchs, Fuchs, & Compton, 2004), and reading disability (RD) prediction (Compton, Fuchs, Fuchs, & Bryant, 2006). Each WIF test consists of a list of 50 words. Students were asked to read the words in one minute, and the total number of correct responses was recorded as the performance score. The 50 words are randomly sampled based on their frequency level. That is, the first 10 words are selected from the most frequent 100 words; the second 10 words are selected from the most frequent 101-200 words, and so on. In this study, two lists of WIF word were assessed every other week. The students were measured on a total of 19 lists of words in Year 1 (I neglected to administer List 2. This missing data point is a mistake that I could not correct later), and 22 lists in Year 2.

CHAPTER III

RESULTS

Year l Measurement Results

Year 1 Measurement results can be found in Table 1 and Figure 1. All students' post-test raw scores were higher than their pre-test scores on both Word ID and Word Attack. On word ID Boy 5 made the most gain in raw score (+31), and Boy 1 made the least gain in raw score (+20). On Word Attack, Boy 1 and Boy 5 gained 4 points; Boy 3, Boy 4, and Boy 5 gained 9 points.

With percentile scores and standard scores, we can compare the target students' scores to the norm. Boy 2 had the lowest pre-test percentiles on both measurements, 2 for Word ID and 9 for Word Attack. That means, his pre-test Word ID score was only equal to or higher than 2% of the children his age, and his pre-test Word Attack score was only equal to or higher than 9% of the children his age. His post-test Word ID percentile was 38, and post-test Word Attack percentile 52. Which means, in post-test, his Word ID score was equal to or higher than 38% of the children his age, and his Word Attack score was above average. His Progress Monitoring (PM) test also showed the ascending trend. The first data point was 3, and he had his highest PM score (15) on list 19. Because of the satisfactory progress he had made across measures, Boy 2 was identified as a responder. Boy 4 and Boy 5 had a similar pattern. Their post-test scores improved substantially, and they were performing close to or above average level after tutoring.

Boy 4 and Boy 5 were also identified as responders. It was very impressive that Boy 2, Boy 4 and Boy 5 all improved more than one standard deviation (for boy 2 and boy 5, close to 2 standard deviations) on their posttest Word Attack measurement.

Boy 3 had been retained before. His pretest performance on Word ID and Word Attack was about average level. His posttest Word ID standard score remained almost the same level, which was still slightly above average compared to the norm. And his posttest standard score on Word Attack improved. His PM scores showed a steady increase. Boy 3 was also considered to be a responder.

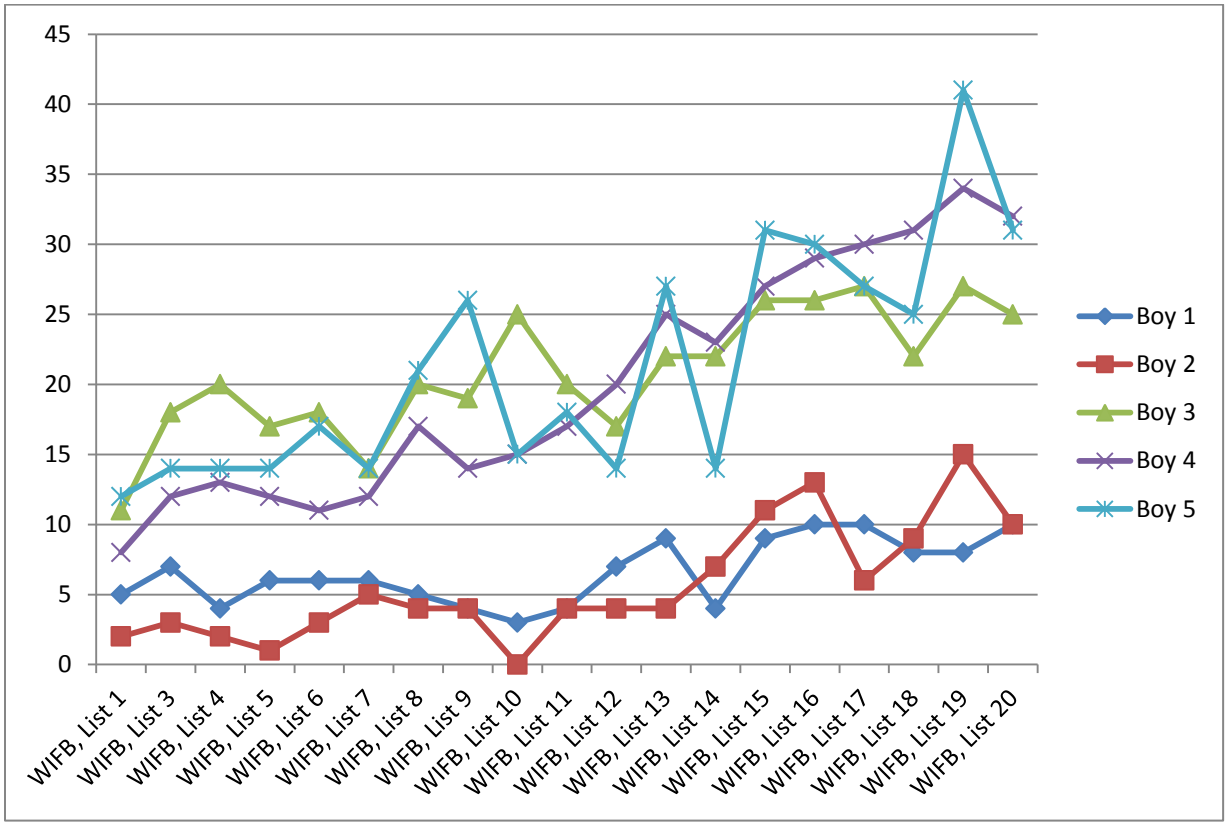
Boy 1's posttest standard scores on both Word ID and Word Attack were lower than pretest standard scores. His PM scores improved slightly, but overall did not show a steep ascending trend as the other participants. Boy 1 was considered non-responsive.

Table 1. Year 1 Word Attach and Word ID

Student	Test	*Pre Raw	*Post Raw	***Gain Raw***	**Pre Per	**Post Per	****Gain Per***	***Pre SS	***Post SS	****Gain SS***	Responsive (Y/N)
Boy 1	Word ID	7	27	***+20***	48	45	***-3***	99	98	***-1***	Y
	Word Attack	4	8	***+4***	64	50	***-14***	105	100	***-5***	Y
Boy 2	Word ID	0	24	***+24***	2	38	***+36***	68	95	***+27***	Y
	Word attack	0	9	***+9***	9	52	***+43***	80	101	***+21***	Y
Boy 3	Word ID	19	41	***+22***	71	70	***-1***	108	108	***0***	Y
	Word Attack	2	11	***+9***	50	60	***+10***	100	104	***+4***	Y
Boy 4	Word ID	15	39	***+24***	65	67	***+2***	106	107	***+1***	Y
	Word Attack	0	9	***+9***	9	44	***+35***	80	98	***+18***	Y
Boy 5	Word ID	15	46	***+31***	65	77	***+12***	106	111	***+5***	Y
	Word Attack	8	12	***+4***	46	62	***+16***	112	104	***-8***	Y

*Pre Raw = Pretest Raw Score; Post Raw = Posttest Raw Score; Gain Raw = Gain Raw Score
**Pre Per = Pretest Percentile Score; Post Per = Posttest Percentile Score; Gain Per = Gain Percentile Score
***Pre SS = Pretest Standard Score; Post SS = Posttest Standard Score; Gain SS = Gain Standard Score

Figure 1. Year 1 Progress Monitoring



*WIFB List 2 is missing because I neglected to administer that list.

Year 2 Measurement Results

Year 2 Measurement results can be found in Table 2 and Figure 2. All students' post-test raw scores were higher than their pre-test raw scores on both Word ID and Word Attack. On word ID Boy 8 made the most gain in raw score (+45), and Boy 6 made the least gain in raw score (+18). On Word Attack, Boy 8 and Girl 2 gained 17 points; Girl 1 gained 12 points; Boy 6 gained 3 points; and Boy 7 gained 2 points.

Boy 7 had the lowest pre-test percentiles on both measurements, 55 for Word ID and 14 for Word Attack. That means, his pre-test Word ID score was equal to or higher than 55% of the children his age, and his pre-test Word Attack score was only equal to or higher than 14% of the

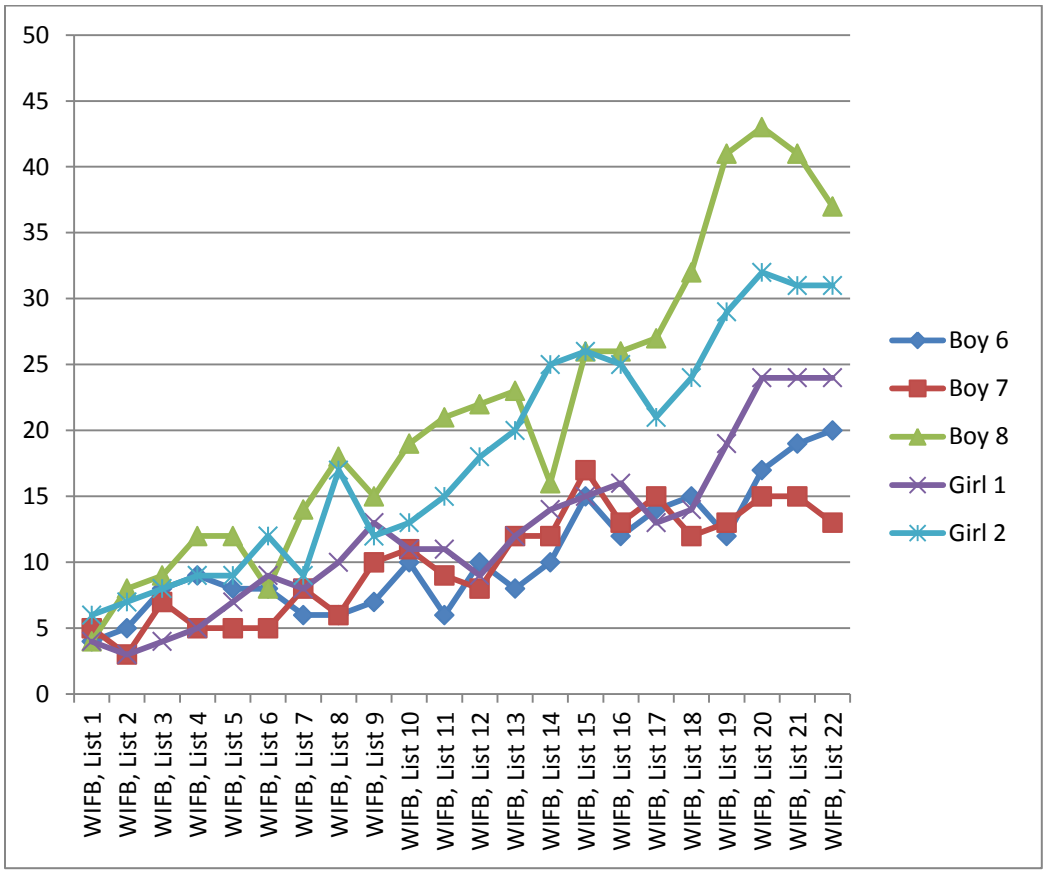
children his age. His post-test Word ID percentile was 53, and post-test Word Attack percentile was 11. This means, in post-test, his Word ID score was equal to or higher than 53% of the children his age, and his Word Attack score was only equal to or higher than 11% of the children his age. His rank on both measures dropped from pretest to posttest. His PM scores also showed the smallest slope. Because he did not make satisfactory progress across measures, Boy 7 was identified as a non-responder. Boy 6 had a similar pattern. He made some minimal progress on Word Attack, but his posttest standard score on Word Attack was still almost one standard deviation below the mean. His posttest Word ID percentile standing dropped 14% from pretest. His PM scores had the second lowest slope. Thus, Boy 6 was also identified as non-responsive to treatment.

Table 2. Year 2 Word Attach and Word ID

Student	Test	*Pre Raw	*Post Raw	***Gain Raw***	**Pre Per	**Post Per	****Gain Per***	***Pre SS	***Post SS	****Gain SS***	Responsive (Y/N)
Boy 6	Word ID	12	30	***+18***	66	52	***-14***	106	101	***-5***	Y
	Word Attack	0	3	***+3***	14	18	***+4***	84	86	***+2***	N
Boy 7	Word ID	7	31	***+24***	55	53	***-2***	102	101	***-1***	Y
	Word attack	0	2	***+2***	14	11	***-3***	84	82	***-2***	N
Boy 8	Word ID	7	52	***+45***	55	85	***+30***	102	115	***+13***	Y
	Word Attack	9	26	***+17***	84	89	***+5***	115	118	***+3***	Y
Girl 1	Word ID	11	38	***+27***	64	65	***+1***	105	106	***+1***	Y
	Word Attack	0	12	***+12***	14	62	***+48***	84	104	***+20***	Y
Girl 2	Word ID	17	39	***+22***	73	67	***-6***	109	107	***-2***	Y
	Word Attack	4	21	***+17***	71	83	***+12***	108	114	***+6***	Y

*Pre Raw = Pretest Raw Score; Post Raw = Posttest Raw Score; Gain Raw = Gain Raw Score
**Pre Per = Pretest Percentile Score; Post Per = Posttest Percentile Score; Gain Per = Gain Percentile Score
***Pre SS = Pretest Standard Score; Post SS = Posttest Standard Score; Gain SS = Gain Standard Score

Figure 2. Year 2 Progress Monitoring



Boy 8, Girl 1, and Girl 2 were all performing above average achieving students on their posttest Word ID and Word attack measures. Their PM scores also showed a clear increasing trend. Girl 1 made an impressive improvement of over one standard deviation on Word Attack. Boy 8's improvement on Word ID was also close to one standard deviation. Boy 8, Girl 1, and Girl 2 were considered to be responders.

CHAPTER IV

DISCUSSION

The majority of the participants responded to the tutoring program and made satisfactory progress. Three of them (Boy 1, Boy 6, and Boy 7) did not respond.

Boy 1 was an African American child who had been retained before. His performance on tutoring and testing was slightly better than Boy 6 and Boy 7. But he still did not make satisfactory progress. Boy 1 was a frequent visitor to the principal's office because of behavior issues. He was suspended for stealing and had the most absences among all Year 1 students. Boy 5 had also been retained before, but he made significant progress across measures. Boy 5 had some attention problems at the beginning of the school year, but he became motivated to learn and enjoyed tutoring. He loved earning points for correct words, and extra points for good behavior.

Boy 6 had a LD diagnosis and received special education at school. It is interesting to note that Boy 2 was in the same school with the same special education placement as Boy 6. But Boy 2 made the most standard score gains among all Year 1 students. Boy 2 always paid full attention during tutoring. He often got frustrated with difficult words, but he kept practicing them. Boy 6 had a very short attention span, and tended to avoid challenging words.

It seems that special education placement does not predict responsiveness. Neither does retaining or pretest performance. What contributes to non-responsiveness? Here are some similarities among the three non-responders:

a. All of them were boys from a low-income family.
b. All of them had attention/behavior problems. Two of them were frequent visitors to the principal's office for misbehavior. All of them had low ratings on the SWAN form from the teachers and the tutor.
c. All of them gave up easily when faced with challenging tasks. For example, when presented with new words, all of them were more likely to say "I don't know" than the other students. The other students usually tried to sound out the words and use the strategies learned in tutoring. These three students did not have the capability to transfer what they learned in tutoring to new words, or same words presented in a different fashion to them.

I was amazed to see the correlation between the SWAN form and responsiveness. All three non-responders had low SWAN ratings. SWAN measures attentiveness and hyperactivity on a continuum. My hypothesis is that the students' learning styles, specifically, their abilities to focus and avoid extraneous distractions, is one important prediction of responsiveness among students with reading difficulties. For students with behavior problems, an effective tutoring program combined with a behavior intervention program may bring more positive results on their academic achievement.

Appendix A. Tutoring Lesson Page Example

Student Name: _____

☐ Completed
☐ 100%
☐ Repeated
MANDATORY

Lesson #8 Word Page – *Fast Dan - Book #2*

Sight Words 

Activity: Word Circle

he	slow	too	man
try	put	even	go
together	by	is	little

Sounds 

Color Sound Card: j

j

g w d r n i f

Decodable Words 

Activity: Sound Sandwiches

fast jump jet job

just got Dan jam

Spelling *A B C*

Activity: Missing Letters

a	b	d	f	g	j	m	o	p	s
t	u	w							

Story Words 

Activity: Find the Story Words

zip won't

even try

REFERENCES

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities, 39* (5), 414-431.

Compton, D.L., Fuchs, D., Fuchs, L.S., & Bryant, J.D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. Journal of Educational Psychology, 98, 394-409.

Fuchs, D., Compton, D., Fuchs, L., Bryant, J., & Davis, N. G. (2008). Making "secondary intervention" work in a three-tier responsiveness-to-intervention model: findings from the first-grade longitudinal reading study of the National Research Center on Learning Disabilities. *Reading and Writing, 21* (4), 413-436.

Fuchs, D., & Fuchs, L. S. (2005). Peer-Assisted Learning Strategies: Promoting word recognition, fluency, and comprehension in young children. *Journal of Special Education, 39*, 34-44.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Responsive to Intervention: What, why and how valid is it? *Reading Research Quarterly, 41* (1), 93-99.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71*, 7-21.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, Pearson Education, Inc.

Swanson, J. M., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J., … McCleary, R. (2001). Categorical and dimensional definitions and evaluations of symptoms of ADHD: the SNAP and SWAN ratings scales. Retrieved from http://www.adhd.net

Woodcock, R. W. (1998). Woodcock Reading Mastery Tests - revised. Circle Pines, MN: American Guidance Service.